Regulation Of Transcription By RNA Polymerase II In *S. pombe*

Thesis

Regulation of transcription by RNA polymerase II in S. pombe

Sneha Gopalan
M.S. Life Sciences

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Stowers Institute for Medical Research,
An Affiliated Research Centre of The Open University
June 2017

Abstract

An understanding of the mechanisms underlying the various stages of transcription is crucial to find solutions to the problems caused by mis-expression of genes that may give rise to a host of human diseases. My thesis research focusses on an analysis of the RNA Polymerase II elongation factor Ell1/Eaf1 in *S. pombe*. Eleven-nineteen lysine-rich in leukemia (ELL) is encoded by a gene involved in translocations with MLL in leukemia and forms a tight complex with ELL-associated factors (EAF). ELL/EAF is an RNA polymerase II elongation factor that in metazoa can assemble into a larger assembly that also includes P-TEFb and other proteins encoded by genes involved in MLL translocations. This larger assembly, sometimes called SEC, binds to a specific "docking site" in the metazoan Mediator complex. Distantly related ELL- and EAF-like genes were identified in *S. pombe* that encode Ell1/Eaf1 and can stimulate Pol II elongation *in vitro*. My thesis addresses two distinct projects, with overlapping motivations:

First, to see whether *S. pombe* might provide a good model for functional studies of the Mediator and ELL/EAF interaction, I carried out a thorough proteomic analysis of *S. pombe* Mediator and defined several new subunits. My results were recently published as part of a collaborative structural analysis of *S. pombe* Mediator.

Second, I used a combination of biochemical, genetic, and genomic approaches to characterize Ell1/Eaf1 function in fission yeast. Using mass spectrometry, I identified an uncharacterized sequence orphan, *SPAC6G9.15c*, that associates with both Ell1 and Eaf1 to form a ternary complex that, based on ChIP-seq localizes at genes with high RNA Pol II occupancy. I also performed an SGA screen for genes that genetically interact with *ell1*, *eaf1*, and *SPAC6G9.15c* and identified a set of overlapping genes that interact with all three, as well as others that interact only with *ell1*.

Dedicated to Mom and Dad….

Acknowledgement

First and foremost, I am grateful to my mentors, Drs. Joan and Ron Conaway, whose expertise, understanding, guidance and support made it possible for me to work on this project. Their honesty, passion for science and attention to detail is highly motivating and has made me a better scientist. I would forever be thankful to them for giving me the freedom to evolve and patiently teaching me values of science and life.

I am thankful to my committee members Peter Baumann and Jerry workman who taught me to critically analyze my work and kept me on track for the timely completion of this thesis. I appreciate the help of various people and Stowers core facility, who contributed and helped in my thesis work; Dana Gibbon, Chris Seidel, Proteomics core and Molecular Biology core. I would like to acknowledge the valuable help and suggestions from my past and present lab members, especially Chieri and Shigeo Sato, Lu Chen and Charles Banks. I thank my colleagues Juston, Soon-Keat, Melvin, Merry, Sara and other lab members for making the lab a fun and wonderful place to work. I am thankful to Leanne for keeping everything from the Open University in order, and to Shelly and Lisa for managing all the deadlines. I would also like to thank my examination committee, Julia Cooper, Linheng Li and Tatjana Piotrowski, for their time and support during my thesis defense.

Finally, I would like to thank my family and friends for all the love and support throughout my life. I am thankful to my parents for always believing in me and making me whatever I am and my little brother, Sachin, who always has ways to make me laugh. Most of all, none of this would have been possible without the encouragement and support of my loving husband, Karthik. He has put up with all my craziness, cheered me up, stood with me through thick and thin, and constantly inspires me to be a better person. I look forward to ‘forever’ with him and my heart that beats outside, my baby Neil.

Table of Contents

List of Figures

List of Tables

Abbreviations used

% (v/v)..... ml per 100ml (volume/volume)
% (w/v)..... grams per 100ml (weight/volume)
°C ..... degrees centigrade
aa ..... amino acid(s)
ATP ..... adenosine 5'-triphoshate
bp..... base pair
BSA ..... bovine serum albumin
Cdk9..... cyclin dependent kinase 9
cDNA..... complimentary deoxyribonucleic acid
ChIP ..... chromatin immunoprecipitation
clonNAT ..... nourseothricin
cpm.....counts per million
CTD ..... carboxy terminal domain
CTP..... cytidine 5' triphosphate
Da ..... Dalton
DEAE ..... diethylaminoethyl
dELL ..... drosophila ELL
DEPC ..... diethylpyrocarbonate
DNA ..... deoxyribonucleic acid
DRB 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole
DSIF..... DRB sensitivity inducing factor
DTT ..... dithiothreitol
EAF..... ELL associated factor
EBP..... ELL binding protein
EDTA ..... diaminoethanetetra-acetic acid
ELL ..... eleven-nineteen lysine rich in leukaemia
EMM ..... Edinburgh minimal medium
GTP ..... guanosine triphosphate
g..... gram
h..... hour
HEPES..... N-[hydroxyethylpiperazine –N '-[2-ethanesulphonic acid]
IGV ..... Integrative Genomics Viewer
IP..... immunoprecipitation
IMP ..... inosine monophosphate
kb..... kilobase
M ..... molar
MDa ..... megadalton

ME..... malt extract
mg..... milligram
min..... minute
ml ..... millilitre
MLL..... mixed-lineage leukaemia or myeloid / lymphoid leukaemia
mM ..... millimolar
mRNA..... messenger RNA
NELF..... negative elongation factor
nm..... nanometre
NP-40..... Nonidet P-40
N-terminal ..... amino terminal
NTP ..... nucleotide triphosphate
OD..... optical density
ORF ..... open reading frame
PAGE..... polyacrylamide gel electrophoresis
PBS..... phosphate buffered saline
PCR..... polymerase chain reaction
PIC..... pre-initiation complex
Pol II..... RNA polymerase II
pre-mRNA..... precursor messenger ribonucleic acid
P-TEFb..... positive elongation factor b
qPCR ..... quantitative polymerase chain reaction
RACE ..... rapid amplification of cDNA ends
RNA..... ribonucleic acid
Rpb1 ..... RNA polymerase B (pol II) subunit 1
RPKM ..... Reads Per Kilobase of transcript, per Million mapped reads
rRNA ..... ribosomal RNA
RT-qPCR..... reverse transcription followed by quantitative PCR
SDS..... sodium dodecyl sulphate
SEC.....Super Elongation Complex
SII ..... factor stimulating RNA pol II - II (TFIIS)
SIII ..... factor stimulating RNA pol II - III (Elongin)
snRNA ..... small nuclear RNA
snRNPs..... small nuclear ribonuclear proteins
S. pombe..... Schizosaccharomyces pombe
Spt5..... suppressor of Ty 5
TBE..... Tris-borate-EDTA buffer
TBP..... TATA binding protein
TE ..... Tris-EDTA buffer

TFIIA..... general transcription factor II A[1]
Tris..... Tris[hydroxymethyl aminomethane]
tRNA ..... transfer RNA
UTP ..... uridine triphosphate
UV..... ultraviolet
V..... volts
YES..... yeast extract with supplements

[1] Similarly for transcription factors TFIIB, TFIID, TFIIE, TFIIF, TFIIH and TFIIS

CHAPTER 1: Introduction

1.1 Regulation of gene expression

Cells need to perform various functions at various times to survive. To do so, cells need to alter the kinds and amounts of proteins they contain in response to growth conditions. The amount of a protein present in a cell can be controlled by regulating many steps of gene expression, including RNA synthesis, the processing of primary transcripts into mature mRNA, transport of mRNA from the nucleus to cytoplasm, translation or turnover of mRNAs, and, ultimately, protein degradation.

Protein coding genes are transcribed into mRNA precursors by RNA polymerase II (Pol II). Synthesis of RNA transcripts by Pol II is the first step in gene expression and is a major point of gene regulation. Transcription can be divided into 3 major steps; initiation, elongation and termination. This introduction will examine the functions and regulation of Pol II, discuss a key transcriptional coregulator called the Mediator complex and the transcription elongation factor ELL, and introduce *S. pombe* as a model system for answering questions about the functions of ELL in cells.

1.2 RNA Pol II transcription

In eukaryotic cells the synthesis of ribonucleic acid (RNA) from the DNA is achieved by the three types of RNA Polymerases present. Pol I is responsible for synthesis of the large ribosomal RNAs (rRNAs), and Pol III for 5S and 5.8S RNAs, tRNAs, and several

other small non-coding RNAs. Pol II is responsible for the synthesis of protein coding mRNAs along with other non-coding RNAs, including snRNAs, which have key roles in RNA splicing, and long non-coding RNAs, miRNAs, and others, many of which regulators of gene expression.

This section of the introduction examines the basic properties of Pol II and its function in basal transcription and the various stages of transcriptional regulation.

RNA Polymerase II

RNA Polymerase II on its own it cannot recognize and bind promoter or unwind DNA template around a transcription start site. However, Pol II can bind to and initiate transcription on single stranded DNA. In addition, once it is initiated, it can transcribe double stranded DNA without help from additional proteins. In cells, however, the DNA is not naked and bound by histone complexes and the various mechanisms that exist for providing the polymerase access to the template are discussed later. Sequence specific factor also help to activate specific genes.

Pol II consists of 12 polypeptides, designated Rpb1 through Rpb12. The largest Pol II subunit, Rpb1, contains a largely unstructured C-terminal domain (CTD) that constitutes a unique feature of Pol II and distinguishes it from all other polymerases. The CTD, composed of tandem heptad repeats, is conserved from fungi to humans, although the number of repeats and the consensus sequence vary in different organisms. For example, vertebrate CTDs contain 52 tandem heptapeptide repeats (most having a consensus $Y^1S^2P^3T^4S^5P^6S^7$), and the fission yeast contains 29 heptads, 24 of which are all-consensus,

whereas the CTD of budding yeast consists of 26 heptads, 19 of which are perfect consensus.

Several kinases have been implicated in phosphorylation of the CTD. The CTD is phosphorylated on Ser5 by a protein kinase, CDK7/Cyclin H, that is associated with the general initiation factor TFIIH. CDK7 also appears to phosphorylate the CTD on Ser7 (Akhtar et al., 2009). CDK9/cyclin T (also called P-TEFb in mammals), phosphorylates the CTD on Ser2 (Meinhart and Cramer, 2004) and is required for phosphorylation of Thr4 (Hsin et al., 2011). Another PIC component, CDK8/Cyclin C, a subunit of the Mediator complex, can also phosphorylate the CTD on both Ser2 and Ser5 *in vitro* (Liao et al., 1995). Serine 5 phosphorylation usually occurs at or near the promoter, while serine 2 phosphorylation is seen primarily on polymerase molecules that have moved away from the promoter region and are engaged in transcript elongation (Komarnitsky et al., 2000).

1.2.A Transcription initiation

Pol II can bind to and initiate transcription on single stranded DNA. In addition, once it has initiated, it can transcribe double stranded DNA without help from additional proteins. On its own, however, it cannot recognize and bind promoter DNA, nor can it unwind the DNA template around the transcription site. Accordingly, promoter-specific transcription by RNA Polymerase II depends on assistance from a set of evolutionarily conserved general initiation factors.

Transcription initiation is a complex process involving the action of a large number of general and gene-specific transcription factors. Five evolutionarily conserved

general initiation factors, TFIIB, TFIID, TFIIE, TFIIF and TFIIH are necessary for promoter-specific transcription initiation to take place *in vitro.* These factors, along with Pol II, are recruited to the promoter to form the pre-initiation complex (PIC) (Conaway and Conaway, 1997; Roeder, 1996). A sixth factor, TFIIA, can stimulate transcription *in vitro* but is not essential.

TFIID, containing the TATA binding protein (TBP), binds to and bends the DNA at the promoter of the gene. TFIIA (Roeder, 1996) stabilizes the interaction between TBP and DNA. TFIIB promotes the selective binding of preformed Pol II/ TFIIF to TBP and in yeast has been shown to direct transcription to begin at a defined start site downstream of the TATA box. TFIIF performs dual roles in assembly of the preinitiation complex by binding to and strongly stabilizing the Pol II-containing intermediate and by supporting recruitment of TFIIE and TFIIH into the complex. Following the full assembly of the preinitiation complex, an ATP dependent DNA helicase activity associated with the XPB subunit of TFIIH helps to unwind the DNA surrounding the transcriptional start site (TSS) to form the open complex prior to initiation by Pol II.

Then initiation takes place, leading to formation of the first phosphodiester bond. The initial transcribing complex is fairly unstable. This instability can result in the formation of short transcripts as the initiation complex disassociates from the DNA, a process known as abortive initiation. Abortive initiation has been proposed to be due at least in part to interference between the growing RNA and the TFIIB B-finger (Kostrewa et al., 2009).

Once transcripts longer than 10-15 have been made, the polymerase goes into productive elongation.

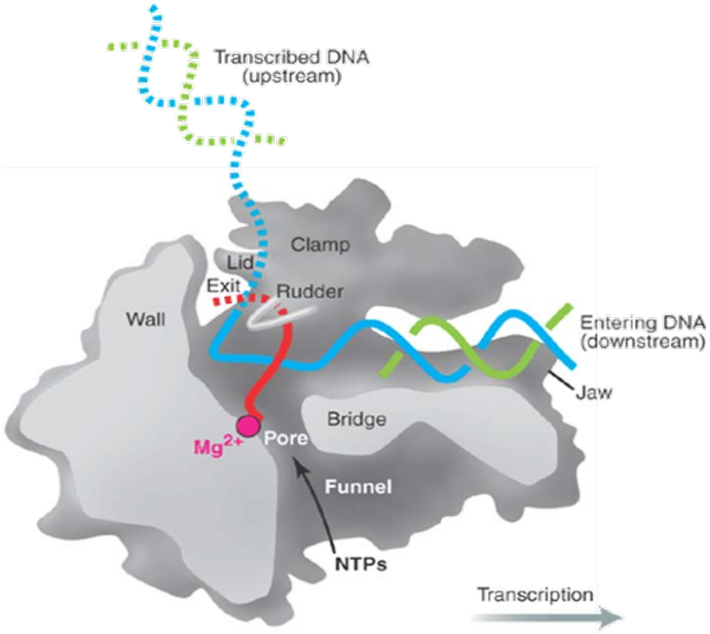


Figure 1.1 A transcribing ternary complex.

Figure is from (Klug, 2001) © Roger Kornberg reprinted with permission.
Cross sectional diagram showing some structural features of the elongating RNA polymerase II. As Pol II transcribes from left to right, the downstream DNA is unwound and the template strand shown in blue is positioned near the 3' end of the growing RNA chain (red) and the active site magnesium ions (pink). Polymerisation of the RNA chain occurs in a template dependent manner as nucleotides access the active site via the "funnel" and "pore". As each nucleotide is added, a change in conformation of the bridge helix causes a translocation of the enzyme with respect to the DNA / RNA positioning the new 3' end of the RNA chain at the active site ready for further nucleotide addition. The RNA chain is separated from the template by the rudder and guided towards the exit channel. During transcript elongation the enzyme remains stably associated with the DNA and RNA; downstream DNA is gripped by a pair of "jaws" (only one of which is shown here) and nucleic acids are also restrained by a "clamp" which swings towards the active site during elongation (Gnatt et al., 2001).

1.2.B Early elongation

CTD phosphorylation: The phosphorylation state of the CTD is critical in determining its activity and plays a central role in coordinating pre-mRNA processing by recruiting many of the enzymes and proteins critical for proper capping, splicing, and

polyadenylation of pre-mRNA, as well as for proper nuclear export and localization of mature mRNAs (Buratowski, 2009; Perales and Bentley, 2009; Yoh et al., 2008).

Promoter escape occurs as the nascent RNA binds stably to Pol II and triggers the release the of TFIIB, after which the ternary transcription complex is referred to as the early elongation complex and energy from ATP hydrolysis is no longer required for transcript elongation (Dvir et al., 2001; Kostrewa et al., 2009). This complex is not as stable as a fully functional ternary complex and is susceptible to transcript slippage or backtracking (until about +30), resulting in transcriptional arrest of Pol II as its active site becomes misaligned with the 3' end of the transcript. The transcription factor TFIIS (SII) can induce Pol II catalyzed transcript cleavage to form a new 3' end properly aligned at the active site, allowing elongation to continue (Pal and Luse, 2003; Wang et al., 2009).

Transcriptionally engaged polymerase is also susceptible to controlled promoter proximal pausing in the region +20 to +40. The duration of pausing or arrest during early elongation can be controlled by multiple transcription elongation factors that influence the elongation competence of Pol II. Pausing is exacerbated by two negatively acting transcription factors, 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB)-sensitive inducing factor (DSIF) and negative elongation factor (NELF), which usually function together to promote Pol II arrest, possibly through binding of NELF to RNA. Release of Pol II from some of the pausing sites to start productive elongation is triggered by positive transcription elongation factor b (P-TEFb)(Peterlin and Price, 2006).

1.2.C Regulation during transcription elongation

Transcriptional elongation is a highly regulated step during gene expression. Efforts to identify activities that stimulate the rate of elongation by RNA Polymerase II led to the discovery of numerous transcription factors, that can interact directly with elongating Polymerase and increase its elongation rate by suppressing transient pausing by the enzyme. These elongation factors include TFIIF and members of the ELL and Elongin families (Shilatifard et al., 1997c; Uptain et al., 1997). DSIF (Spt4/5) can also stimulate elongation via suppression of transient pausing under some conditions. Additional elongation factors, including TFIIS and the PAF1c can stimulate elongation by other mechanisms. Below I summarize properties of IIF, Elongin, DSIF, TFIIS, and PAF1c. ELL will be discussed in section 4.

TFIIF participates in suppression of arrest of very early Pol II elongation intermediates at least in part by functioning as an adaptor that links TFIIE and TFIIH to the transcribing Polymerase. It also has an elongation activity capable of increasing the rate of synthesis of the first few phosphodiester bonds of nascent transcripts, ensuring that growing transcripts reach a sufficient length (>6 to 8 nucleotides) to be relatively resistant to abortion (Yan et al., 1999). TFIIF dissociates from early RNA Polymerase II elongation complexes shortly after Polymerase has synthesized 9 or 10 nucleotide long transcripts (Yudkovsky et al., 2000).

Elongin was originally purified from rat liver nuclei by its ability to stimulate the rate of elongation by RNA polymerase II *in vitro* (Bradsher et al., 1993). Elongin is a heterotrimeric protein composed of a transcriptionally active Elongin A subunit and two

smaller Elongin B and Elongin C subunits. Elongin A can also act as the substrate recognition subunit of a Cullin-RING E3 ubiquitin ligase that has been shown to target Pol II stalled at sites of DNA damage (Harreman et al., 2009; Ribar et al., 2007; Yasukawa et al., 2008).

ELL and its function will be discussed in the fourth section of this introduction.

Polymerase associated factor complex (Paf1C) is another transcription elongation factor that coordinates the movement of Pol II through chromatin and the cotranscriptional processing and fate of nascent transcripts. It is composed of the proteins Paf1, Ctr9, Cdc73, Rtf1, and Leo1 (Wade et al., 1996). The Paf1C has been linked to a large and growing list of transcription related processes including: communication with transcriptional activators; recruitment and activation of histone modification factors (H2B ubiquitylation, H3K4/h3K36 methylation); facilitation of elongation on chromatin templates; and the recruitment of 3' end-processing factors necessary for accurate termination of transcription. It also acts as a regulator of promoter proximal pausing (Chen et al., 2016).

1.2.D Termination and RNA processing

Pol II transcripts are synthesized as precursors that are processed by capping, splicing, and 3’-end processing, which results in formation of 3’- polyadenylated transcripts (most mRNAs, some non-coding RNAs) or non-polyadenylated transcripts (Richard and Manley, 2009). In eukaryotes, the terminator signals are recognized by protein factors (cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulation factor

(CstF)) which trigger the termination process. Once signals directing polyadenylation are transcribed, these factors recruit other proteins to the site to cleave the transcript, freeing the mRNA from the transcription complex, and add a string of about 200 A-residues to the 3' end of the mRNA in a process known as polyadenylation. During these processing steps, the RNA polymerase continues to transcribe for several hundred to a few thousand bases and eventually dissociates from the DNA and downstream transcript.

pre-mRNA transcripts synthesized are processed by addition of a 7-methylguanosine cap at the 5’ end and splicing out of introns prior to transcription termination and export of the transcript from the nucleus.

RNA capping of nascent transcripts takes place shortly after transcription initiation and is crucial for mRNA stability, splicing, polyadenylation, export and translation efficiency. In metazoans, a bifunctional capping enzyme with RNA triphosphatase and RNA guanylyltransferase activities binds to the phosphorylated CTD through the guanylyltransferase domain. This domain has two binding sites for phosphorylated CTD: one specific for the Ser2 phosphorylated CTD and the other, an allosteric activator site, for the Ser5 phosphorylated CTD (Ho and Shuman, 1999). In fission yeast it was shown that that recruitment of capping enzymes is the sole essential function of Ser5 as substitution of all Ser5 residues with Ala is lethal in *S. pombe*, but viability can be restored simply by tethering capping enzymes to the CTD (Schwer and Shuman, 2011). Capping enzyme can also bind to Spt5 (a subunit of DSIF) and Cdk9 (a component of P-TEFb) suggesting coordination between capping and control of early elongation, which may prevent productive elongation of uncapped transcripts (Pei et al., 2003; Rasmussen and Lis, 1993).

Splicing of the pre-mRNA containing introns and transcription is coupled in cells. The spliceosome complex consists of snRNPs (small nuclear ribonucleoproteins) and other proteins (non-snRNPs) including members of the SR family. Hyperphosphorylated Pol II coimmunoprecipitates with both snRNPs and SR proteins suggesting a physical interaction between Pol II and the spliceosome (Kim et al., 1997).

3' end processing of RNA Pol II transcripts is also mediated by CTD phosphorylation. Polyadenylation of the 3' end is a simple two-step reaction consisting of an endonucleolytic cleavage catalyzed by CPSF, followed by poly(A) tail synthesis on the cleaved transcript by polyadenylate polymerase, which binds CPSF. Polyadenylation serves to protect the 3' end of the transcript from degradation and aids RNA export and translation. Impairment of Ser2 phosphorylation by deletion of CTK1 in yeast (Ahn et al., 2004) or treatment with the P-TEFb inhibitor flavopiridol in metazoan cells (Ni et al., 2004) impairs recruitment of processing factors at the 3' ends of genes and subsequent polyadenylation. The Polymerase continues to transcribe even after transcript cleavage until directed to terminate, but the resulting RNA is unmodified and quickly degraded.

1.3 The Mediator complex

A pivotal role is played by the multi-protein Mediator, which acts as a coregulatory complex for RNA Pol II (Myers and Kornberg, 2000). First isolated from *Saccharomyces cerevisiae* (Kelleher et al., 1990), Mediator is essential for the regulated basal transcription of all genes by Pol II (Takagi and Kornberg, 2006; Thompson and Young, 1995). Subsequent investigation to understand the mechanism of action of

Mediator has revealed that Mediator promotes activation of Pol II transcription via direct interactions with both DNA binding transcription factors and the Pol II preinitiation complex. Further studies have identified an array of Mediator surfaces capable of binding specifically to the transcription activation domains (TADs) of a large number of DNA binding transcription factors and to Pol II and several of the general factors (Conaway and Conaway, 2013). Mediator plays a role (i) in overcoming the activities of factors that negatively regulate elongation (Jishage et al., 2012; Malik et al., 2007), (ii) in recruiting Pol II transcription elongation factors and pre-mRNA processing factors (Donner et al., 2010; Takahashi et al., 2011) and (iii) in controlling phosphorylation of the heptapeptide repeats in the Pol II C-terminal domain (CTD) (Boeing et al., 2010; Kim et al., 1994).

1.3.A Mediator structure and function

Whereas Mediator has some 30 subunits in humans (Sato et al., 2004), there are only 20 subunits in *S. cerevisiae* (Kim et al., 1994). Biochemical studies and 3D structure of budding yeast Mediator reveals that it consists of 3 submodules, referred to as the "Head", "Middle", and "Tail". A fraction of the mediator in cells can also associate with an additional Kinase module having 4 subunits (MED12, MED13, Cyclin dependent kinase CDK8 and Cyclin C) (Asturias et al., 1999; Dotson et al., 2000). Another fraction of Mediator that is free of the Kinase module can associate with Pol II to form what is called the "holoenzyme". It has been suggested that the Kinase module must be displaced for the Mediator to interact with the Pol II (Elmlund et al., 2006; Knuesel et al., 2009). Proteomic analysis helped to define a complete set of human Mediator subunits and

found that the metazoan holoenzyme includes an additional subunit known as MED26 (Sato et al., 2004). A small population of mammalian mediator also contains both Kinase module and Med26.

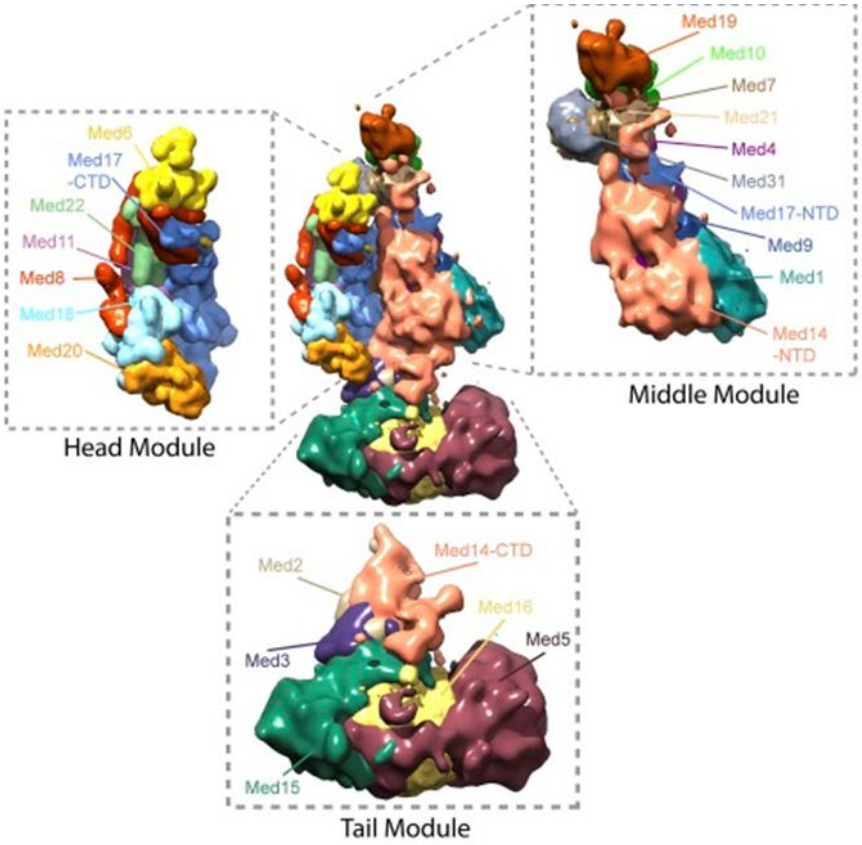


Figure 1.2 Mediator complex architecture.

Figure is from (Robinson et al., 2015) Copyright © 2015, Robinson et al.

S. cerevisiae Mediator subunit localization density map colored by individual subunit.

Mediator can act both as a transcription activator and repressor. Experiments in both yeast and humans suggested that Mediator containing Kinase module contributes to the repression of Pol II transcription (Green and Johnson, 2004). The Mediator lacking the Kinase but containing the MED26 acts as an activator of transcription in mammalian cells and support activation of transcription *in vitro*. More recent studies, however,

reveal that cdk8 kinase activity is required for activation of a number of genes (Donner et al., 2007; Furumoto et al., 2007).

1.3.B Role of Mediator in transcription elongation

An early report showing the recruitment of Mediator to *Drosophila* heat shock genes only upon heat shock coinciding with the release of constitutively paused Pol II (Park et al., 2001), suggested that Mediator may have roles in transcription elongation. It was later observed in mouse ES cells that deletion of Med 23 interfered with the recruitment of Mediator and resulted in a failure to release paused pol II (Wang et al., 2005). Also, Mediator subunits can be detected by ChIP not only at promoters but also within the transcribed regions of genes in both yeast and human cells (Takahashi et al., 2011; Zhu et al., 2006).

A role of Mediator in the regulation of transcript elongation has emerged as it helps in overcoming promoter proximal pausing caused by factors like DSIF and NELF (Wada et al., 1998; Yamaguchi et al., 1999), by recruiting the elongation factors to the genes and controlling the phosphorylation state of the C- terminal domain of the largest Pol II subunit (Donner et al., 2010; Takahashi et al., 2011).

Studies suggest an interaction between P-TEFb and the Kinase module, and depletion of CDK8 reduced both P-TEFb recruitment and Pol II CTD phosphorylation (Donner et al., 2010). Mediator subunit MED23 also regulates basal transcription *in vivo* via an interaction with P-TEFb (Wang et al., 2013).

In metazoa, the MED26 subunit of Mediator acts as a docking site for elongation factors and super elongation complexes (SEC) (Takahashi et al., 2011). N-terminal domain (NTD) of Med 26 can bind directly to the SEC *via* its EAF subunit, hence Mediator supports activator dependent recruitment of ELL-EAF-containing complexes, including SEC, to promoter DNA *in vitro* (Takahashi et al., 2011). TFIIS, Elongin A and IWS1 all have a domain closely resembling the Med 26 NTD and have roles in Pol II elongation. MED26 NTD also interacts with general transcription factor TFIID during initiation and hence may contribute to the switch of Pol II from initiation to productive elongation (Takahashi et al., 2011).

1.4 ELL protein

ELL stimulates transcription elongation *in vitro* and has been shown to interact physically with Pol II (Shilatifard et al., 1997c). ELL was first purified from rat liver nuclear extracts as an ~80kDa single polypeptide that can increase the rate of transcription by Pol II *in vitro* by suppressing transient pausing by Polymerase at multiple places along the DNA (Shilatifard et al., 1996). Two homologues of ELL have since been identified in humans, ELL2 (Shilatifard et al., 1997b) and ELL3 (Miller et al., 2000).

How ELL stimulates elongation by Pol II has not been unequivocally determined. Based in part on the observation that ELL and other transcription factors with similar activities, including Elongin and TFIIF, can inhibit TFIIS induced nascent transcript cleavage by non-arrested RNA Polymerase II elongation intermediates (Elmendorf et al.,

2001), it has been proposed that such factors suppress pausing by preventing displacement of the 3'-end of the nascent transcript from the polymerase catalytic site.

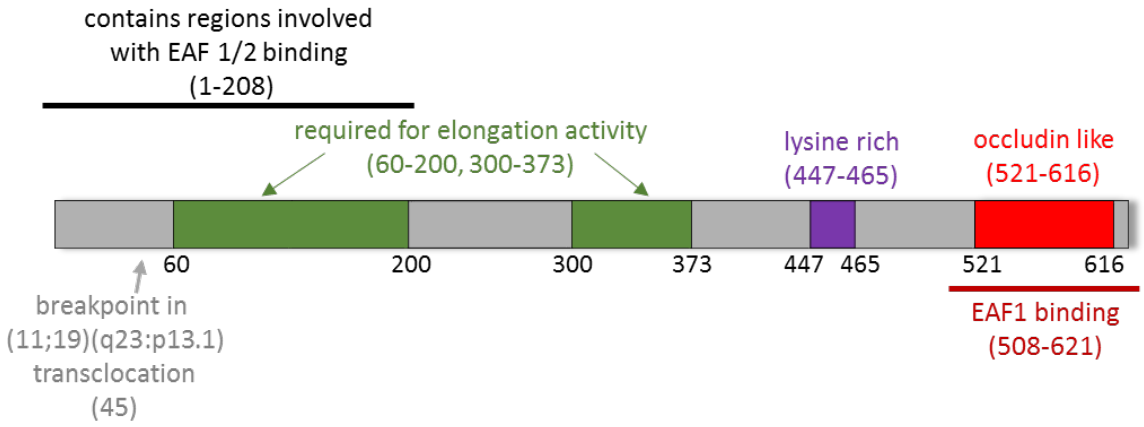


Figure 1.3 The human ELL protein.

The locations of the regions indicated in the figure are described: ELL full length gene, lysine rich region and translocation breakpoint (Thirman et al., 1994), regions required for stimulation of transcription elongation (Shilatifard et al., 1997c), occludin like domain (Shilatifard, 1998a), regions involved in EAF1 and EAF2 binding (Simone et al., 2003).

Structure-function studies have defined a number of functional domains in ELL (Figure 1.3), including regions needed for stimulation of elongation and a small N-terminal region that is dispensable for elongation activity but can interfere with promoter-specific transcription initiation by Pol II (Shilatifard et al., 1997a). In humans and other metazoan organisms, ELL protein has a C-terminal domain similar to occludin, a membrane protein which localizes to tight junctions. The occludin-like domain is not required for ELL elongation activity *in vitro* (Shilatifard, 1998b). Other regions have been mapped including regions required for binding to ELL-associated factors (EAFs) and

regions that have found to inhibit transcription initiation or stimulate transcription elongation.

The earliest information about *in vivo* functions of ELL came from studies in *Drosophila*. The *Drosophila* gene encoding the ELL homolog *Suppressor of Triplo-lethal* (*Su*(*Tpl*)) is essential for viability in flies. Loss of ELL gives rise to embryonic segmentation defects (Eissenberg et al., 2002). Some alleles of *Su(Tpl*) suppress lethality resulting from overexpression of the *Tpl* gene, perhaps by impairing synthesis of *Tpl* mRNA (Eissenberg et al., 2002). *Drosophila* ELL colocalizes with Pol II at transcriptionally active sites on polytene chromosomes, and it was proposed that mutations in Su(Tpl) might preferentially affect synthesis of some long transcripts (Gerber et al., 2001). Although the Pol II CTD is not required for Pol II binding or function of ELL *in vitro*, reduced phosphorylation after knockdown of the cdk9 subunit of P-TEFb in *Drosophila* resulted in decreased binding of dELL, suggesting CTD phosphorylation could contribute to ELL recruitment (Eissenberg et al., 2007). Like the Cdk9 subunit of P-TEFb and the Spt5 subunit of DSIF, a substantial amount of dELL colocalizes with phosphorylated Pol II on polytene chromosomes at transcriptionally active puff sites and relocalizes to heat shock genes upon heat shock (Thirman et al., 1997), consistent with the idea that ELL functions as elongation factor *in vivo*.

In *C. elegans* ELL and EAF proteins play important roles in fertility, survival and body size regulation, in part by regulating the cuticle synthesis, suggesting ELL-EAF may help in the regulation if the extracellular matrix components (Cai et al., 2011). In *Xenopus laevis*, EAF2 protein functions during eye development to activate transcription of the gene encoding the essential Rx homeodomain transcription factor (Maurus et al., 2005).

ELL2 also plays an important role in immunoglobin secretion in plasma cells. It directs efficient alternative mRNA processing, influencing both proximal poly(A) site choice and exon skipping with genes encoding immunoglobulin heavy chain complex (IgH) (Martincic et al., 2009). Reducing ELL2 by siRNA, which reduced processing to the secretion-specific poly(A) site, also influenced the methylations of histone H3K4 and H3K79 on the IgH gene and impacted positive transcription factor b (P-TEFb), Ser-2 carboxyl-terminal phosphorylation, and polyadenylation factor additions to RNA Polymerase II (Milcarek et al., 2011).

1.4.A The MLL-ELL chimera in leukemia

The human ELL gene was initially identified as a gene that undergoes fusion to the MLL gene in t(11;19)(q23;p13.1) translocations in acute myeloid leukemia (Thirman et al., 1994). It was named ELL for eleven-nineteen lysine-rich in leukemia. A highly basic, lysine-rich motif of the ELL protein is homologous to similar regions of several proteins, including the DNA-binding domain of poly(ADP-ribose) polymerase.

Chromosomal translocations involving the MLL and ELL genes generate a chimeric MLL-ELL gene that encodes a fusion protein that contains the entire ELL elongation activation domain and the occludin-like domain, but that lacks the first 45 N-terminal amino acids of the ELL protein. ELL can inhibit transcription initiation *in vitro* as it contains a Pol II interacting domain capable of negatively regulating polymerase activity in promoter-specific manner. In MLL-ELL translocation a portion of this functional domain is deleted, and ELL mutants lacking sequences deleted by the translocation bind RNA

polymerase II and are fully active in elongation, but fail to inhibit initiation (Shilatifard et al., 1997c)

MLL family of proteins are H3K4 methyltransferases. The H3K4 methyl mark is widely associated with gene activation. MLL proteins have N-terminal AT-hooks and a catalytic SET domain at the C-terminal (Tkachuk et al., 1992). Apart from H3 K4 methyltransferases in mammals, several (hSET1 MLL1, MLL2) have been found in large complexes. While sharing some subunits (e.g., Ash2L and WDR5), these complexes nevertheless contain unique sets of proteins that suggest nonoverlapping functions (Vedadi et al., 2017). The MLL gene is a recurrent site of genetic rearrangements associated with childhood hematological malignancies (Mohan et al., 2010). Translocations involving MLL can generate in-frame fusions of MLL with over 50 different partner genes, among which are the MLL fusion partners, AF9 family members AF9 and ENL, and the AF4/FMR2 family members AFF1 and AFF4. The critical feature of these chromosomal rearrangements is the generation of a chimeric transcript consisting of 5' MLL and 3' sequences of the gene on the partner chromosome (Luo et al., 2012). No consistent homologies or motifs among the partner gene sequences have been identified that might explain how their fusion to MLL results in leukemia. EAF1 contains a limited region of homology with the AF4, LAF4, and AF5q31 proteins that fuse to MLL in 11q23 chromosome translocations. This domain is rich in serine, aspartic acid, and glutamic acid residues and has been shown to activate transcription.

The MLL-ELL chimera causes immortalization when introduced into hematopoietic progenitor cells, which in turn causes acute myeloid leukemia when the cells are

transplanted into mice (Lavau et al., 2000). The chimeric MLL-ELL protein that results from the (11;19)(q23;p13.1) translocation contains the amino-terminal region of MLL, including its AT hooks, methyltransferase domain, and repression domain, fused to amino acids 46 to 621 of ELL, including its elongation domain, lysine-rich region, and occludin homology domain. Structure -function analyses of MLL-ELL chimera indicate that although the ELL Occludin-like CTD is sufficient for immortalization of hematopoietic precursors *in vitro*, both the ELL elongation activation domain near the N-terminus of the protein and Occludin-like C-terminus appear to contribute to oncogenesis induced by the MLL-ELL fusion protein in mice (DiMartino et al., 2000; Luo et al., 2001).

1.4.B ELL-associated Factors (EAF)

Two homologous proteins, EAF1 and EAF2, have been identified that interact directly with human ELL and positively regulate ELL elongation activity *in vitro* (Kong et al., 2005; Simone et al., 2001; Simone et al., 2003). EAF 1 and 2 bind directly to ELL and are positive regulators of ELL elongation activity (Kong et al., 2005).

Binding to the EAF proteins protein is achieved through regions in both N and C termini of ELL (for EAF1) or through the C terminus only (for EAF2). Upregulation of ELL elongation activity *in vitro* requires only the N terminal region of EAF1 (Conaway lab (Charles Banks) unpublished data). In addition to stimulating ELL activity, the C terminus of EAF1 contains an acidic transactivation domain which, when fused to a GAL4 DNA binding domain, will stimulate transcription from template containing GAL4 binding sites in whole cell extracts (Simone et al., 2001).

In mammalian cells, EAF1, EAF2, and ELL are colocalized in Cajal bodies, nuclear structures that are enriched in factors involved in transcription and mRNA processing (Gall et al., 1999).

1.4.C ELL as a part of SEC

A number of labs found the association of ELL family members with other elongation factors. Using sequential affinity-purification it was shown that human transcription factors/coactivators AFF4, ENL, AF9, and elongation factor ELL2 are components of the Tat-P-TEFb complex (He et al., 2010). A related study found that HIV-1 Tat assembles a multifunctional transcription elongation complex consisting of core active P-TEFb, MLL-fusion partners involved in leukemia (AF9, AFF4, AFF1, ENL, and ELL), and PAF1 complex (Sobhian et al., 2010). Detailed biochemical purifications of some of the most frequently occurring MLL chimaeras resulted in the isolation of the super elongation complex (SEC) (Lin et al., 2010). The SEC contains the ELL family members ELL1, ELL2 and ELL3, along with ELL associated factors EAF1 or EAF2, the MLL translocation partners AF4/FMR2 family member 1 (AFF1; also known as AF4), AFF4, eleven-nineteen leukemia (ENL) and ALL1-fused gene from chromosome 9 (AF9), and the Pol II elongation factor P-TEFb (Lin et al., 2010). The SEC was also independently identified as a complex that binds Mediator through the Mediator subunit Med26 (Takahashi et al., 2011).

P-TEFb is a kinase composed of cyclin-dependent kinase 9 (CDK9) as a catalytic subunit and cyclin T1 (CYCT1) or CYCT2 as a regulatory subunit. P-TEFb phosphorylates

the Pol II CTD, NELF and the Spt5 subunit of DSIF and hence stimulates the dissociation of NELF from Pol II. NELF dissociation allows Pol II to enter the phase of productive elongation (Lis et al., 2000; Marshall et al., 1996; Peterlin and Price, 2006; Wei et al., 1998). Most P-TEFb in cells is sequestered in an inactive complex with the 7SK snRNA and the proteins HEXIM1 or HEXIM2, LARP7 and MEPCE (Zhou et al., 2012). Upon receipt of cellular signals, P-TEFb is released from the 7SK snRNP complex, allowing it to interact with bromodomain-containing protein 4 (BRD4) or to incorporate into the SEC.

AFF4 is required for SEC stability and proper transcription by poised RNA polymerase II in metazoans. Knockdown of AFF4 in leukemic cells leads to a reduction in MLL chimera target gene expression, suggesting that AFF4/SEC could be a key regulator in the pathogenesis of leukemia through many of the MLL partners (Lin et al., 2010). AFF4 is known to act as a scaffolding protein which uses separate domains to bind different SEC subunits. It promotes the interaction between ELL2 and P-TEFb and helps maintain the integrity of SEC. AFF1 has been shown to interact with AFF4 (Yokoyama et al., 2010), though it is not clear whether the interaction between these two paralogous proteins can exist in a single SEC complex. AF9, AF10, and ENL have been shown to act as positive regulators of P-TEFb activity. They also interact directly with the histone methyltransferase Dot1, leading to the suggestion that Dot1-mediated methylation of H3K79 could be central to leukemogenesis in patients with MLL translocations (Bitoun et al., 2007; Mueller et al., 2007; Mueller et al., 2009). The AF9 YEATS domain binds strongly to histone H3K9 acetylation and, to a lesser extent, H3K27 and H3K18 acetylation, which is important for the chromatin recruitment of the H3K79 methyltransferase DOT1L.

However, how H3K79 methylation by Dot1 could lead to gene activation is not known. Dot1 neither resides in nor associate with SEC.

Reports suggest that ELL facilitates Pol II pause site entry and release, as loss of ELL destabilizes the pre-initiation complexes and results in disruption of early elongation and promoter proximal chromatin structure before recruitment of AFF4 and other super elongation complex components. These changes result in significantly reduced transcriptional activation of rapidly induced genes (Byun et al., 2012).

Recently, in both *D. melanogaster* and mammals ELL was found to be part of a second complex named LEC (little elongation complex), which contains ICE1 (interacts with the C terminus ELL subunit 1; also known as KIAA0947) and ICE2 (also known as NARG2) (Smith et al., 2011; Takahashi et al., 2011). LEC plays a role in the expression of Pol II- transcribed snRNA genes (Hu et al., 2013; Smith et al., 2011; Takahashi et al., 2015). LEC also has been implicated in the resumption of transcription after DNA repair (Mourgues et al., 2013a). ELL, EAF1, ICE, and ICE2 were all identified in a proteomic screen for proteins that interact with the transcription / DNA repair factor TFIIH. ELL was shown to be recruited to UV-damaged chromatin dependent on the TFIIH-associated protein kinase Cdk7. In this study, ELL depletion was seen to hinder Pol II transcription resumption after lesion removal and DNA gap filling and increase Pol II retention to the chromatin.

1.4.D Recruitment of SEC

SEC can be recruited to genes it regulates by a variety of mechanisms. In HIV-1-infected cells, Tat activates HIV-1 transcription by recruiting P-TEFb, ELL/EAF, and SEC family members through interactions with the TAR sequence in the 5' end of the nascent HIV-1 transcript (Peterlin and Price, 2006; Sobhian et al., 2010). Tat also promotes SEC formation, which in turn stabilizes ELL2, an otherwise short-lived protein rapidly degraded by the proteasome. MLL–SEC is formed following translocation of MLL to any of the genes encoding SEC components. DNA binding domain of MLL help deliver the SEC to the MLL-target genes and hence achieve aberrant activation of MLL chimaera target genes through misregulation of transcriptional elongation checkpoint control (TECC) (Mueller et al., 2009).

In Human cells, MED26 functions in part by recruiting ELL/EAF- and P-TEFb-containing complexes, including the SEC to promoters of a subset of genes via a direct interaction with the N-terminal domain (NTD) of MED26. EAF1 and EAF2 can bind directly to the NTD of Med26. The MED26 NTD also binds TFIID, and TFIID and elongation complexes interact with MED26 through overlapping binding sites suggesting that MED26 NTD may function as a molecular switch that contributes to the transition of Pol II into productive elongation (Takahashi et al., 2011).

The YEATS domain of ENL/AF9 can also help target SEC to Pol II on chromatin through contacting the human Polymerase-Associated Factor complex (PAFc) complex (He et al., 2011).

The metazoan-specific Mediator subunit MED23 also helps in the recruitment of elongation factor P-TEFb, via an interaction with its CDK9 subunit (Wang et al., 2013). Whether this interaction can help in recruitment of SEC to genes is not known.

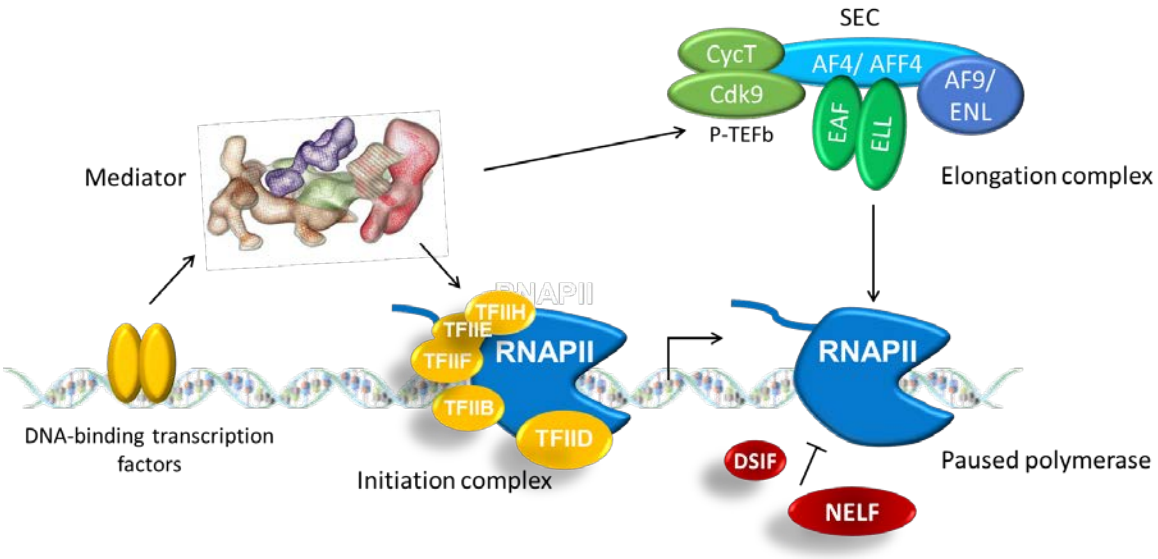


Figure 1.4 Cartoon showing mechanisms of Transcriptional Regulation by SEC.

The metazoan Super Elongation Complex can be recruited to genes by the Mediator through interaction with its subunits Med 26and Med23. SEC can help the paused polymerase to go into productive elongation

1.4.E Other functions of ELL

ELL can specifically interact with the p53 tumor suppressor protein by mechanisms that remain unclear. ELL acts as a negative regulator of p53 in transcription. p53 also inhibits the transcription elongation activity of ELL, suggesting the existence of a mutually inhibitory interaction between p53 and ELL. Elevated levels of ELL in cells resulted in the inhibition of p53-dependent induction of endogenous p21 and substantially protected cells from p53-mediated apoptosis that is induced by genotoxic

stress (Shinobu et al., 1999). ELL may also serve as a transcriptional factor to directly induce transcription of the thrombospondin-1 (TSP-1) gene, which encodes for an anti-angiogenic protein (Zhou et al., 2009).

ELL has been proposed to also function as an E3 ubiquitin ligase and target c-Myc for proteasomal degradation. UbcH8 serves as a ubiquitin-conjugating enzyme in this pathway. ELL-mediated c-Myc degradation inhibits c-Myc-dependent transcriptional activity and cell proliferation, and also suppresses c-Myc-dependent xenograft tumor growth (Chen et al., 2016).

1.5 *S. pombe* as a model system

S. pombe, often known simply as "fission yeast," is an ascomycete yeast. As eukaryotes, these yeasts can be used to study processes that are conserved from yeast to humans but absent from bacteria, such as organelle biogenesis and cytoskeletal organization, or to study mechanisms such as transcription, translation, and DNA replication, in which the eukaryotic components and processes are significantly different from those of their bacterial counterparts (Hoffman et al., 2015). Both *Schizosaccharomyces pombe* and its distant cousin *Saccharomyces cerevisiae* have been studied extensively and have led to the discovery of various genes involved in fundamental mechanisms of cellular processes. *S. pombe* and *S. cerevisiae* share substantial gene content, with ~75% of *S. pombe* genes having one or more recognizable orthologs in *S. cerevisiae*. Nevertheless, these species are separated by as much as 1

billion years of evolution (Hedges, 2002) and therefore, not surprisingly, the biology of these two organisms is dissimilar in several important ways. For example, *S. pombe* cells divide by medial fission, a process analogous to the division of cells in many metazoan organisms, whereas *S. cerevisiae* cells divide by budding. Furthermore, in *S. pombe*, mating and meiosis are tightly coupled, such that only the zygote is ever a diploid, and even here only transiently. By contrast, in *S. cerevisisae*, mating and meiosis are not coupled, and this organism prefers the diploid state. Species-specific gene gains and losses are also apparent. Specifically, genes required for functional complexes involved in pre-mRNA splicing, RNAi-mediated heterochromatin silencing, and signalosome function are present in *S. pombe* (and other metazoan organisms, including humans) but lost in *S. cerevisiae* (Aravind et al., 2000; Clarke, 1990). The structure of the centromeres in *S. pombe* is also considerably more complex and metazoan-like in comparison to the relatively simple centromeres of *S. cerevisiae* (Clarke, 1990). Genome-wide microarray, protein localization, and proteomic analyses suggest moderate conservation between the expression, accumulation, and subcellular localization of orthologous gene products and proteins in these two yeasts (Matsuyama et al., 2006). Since the divergence of the two species approximately 350 million years ago, *S. pombe* appears to have evolved less rapidly than *S. cerevisiae* so that it retains more characteristics of the common ancient yeast ancestor, causing it to share more features with metazoan cells. *S. pombe* also uses some enzyme systems or components of systems that are not present in *S. cerevisiae*, for example RNA interference and parts of the spliceosome.

Basic transcription machinery is conserved from yeast to humans. Previous data from our lab identified genes in *S. pombe* encoding proteins similar to ELL and EAF, which

are not present in *S. cerevisiae* (Banks et al., 2007). Like their mammalian counterpart, SpELL-SpEAF interact with each other and stimulates RNA Polymerase II transcription elongation and pyrophosphorolysis *in vitro*. Hence fission yeast serves as a simpler system to genetically, biochemically and genomically study molecular mechanism elongation complexes.

1.6 Scope of this thesis

1.6.A Using fission yeast to explore the role of mediator and ELL complex in cells

The Mediator complex isolated from the fission yeast *Schizosaccharomyces pombe* contains orthologs of the Head and Middle subunits of *S. cerevisiae*. However, subunits corresponding to the Tail module (MED2, MED3, MED5, MED15 and MED 16) were not found to be associated with this complex in *S. pombe* (Beve et al., 2005a; Spahr et al., 2001). The Tail module is important for the interaction of *S. cerevisiae* and metazoan Mediator with the DNA binding transcriptional activators (Myers et al., 1999). Even though *S. pombe* Mediator appears to lack the Tail domain, the genome of *S. pombe* encodes a putative MED15 ortholog (Linder et al., 2008a). *S. pombe* MED15 is reported to exist in a protein complex along with Hrp1, a CHD1 family ATP dependent chromatin remodeling protein (Khorosjutina et al., 2010). In addition, cryo-EM studies of the purified *S. pombe* Mediator complex were consistent with the idea that it may include only the Head and Middle domains (Elmlund et al., 2006).

ELL and EAF orthologs were not identified in fungi until recently. Previous data from our lab identified genes in *S. pombe* similar to ELL and EAF (*ell1+* and *eaf1+*), which are not present in *S. cerevisiae*. Ell1 and Eaf1 from *S. pombe* also interact with each other to form a heterodimer and can positively regulate transcription by Pol II (Banks et al., 2007). The mechanism of their recruitment in fungi has not been studied, and whether Mediator plays a role in it remains to be determined. Like many other elongation factors in yeast, deletion of the *ell1+* gene in cells makes them sensitive to the drug 6-azauracil (Banks et al., 2007).

Fission yeast is a valuable model system, having the advantage that it is easy to manipulate genetically and hence provides an excellent model for studying the regulation by Ell1 and Eaf1 *in vivo* and for exploring the potential roles of Mediator in orchestrating their activity. It has homologues of P-TEFb, raising a possibility that it may have complexes similar to mammalian super elongation complex (SEC).

1.6.B Questions to be addressed

Motivated by evidence that Mediator helps to regulate ELL function in metazoa, I initiated my thesis research with the goal of exploring potential connections between the Mediator complex and the functioning of Ell1 and Eaf1 in *S. pombe*. An essential first step in addressing this question was to understand the composition of *S. pombe* Mediator, which had not yet been established. Accordingly, this thesis begins by defining the Mediator complex of fission yeast by MudPIT proteomic analysis, in Chapter 2. I find a previously uncharacterized Tail module subunit that associate with the *S. pombe*

Mediator. Also, other subunits were identified that were not shown to be components of *S. pombe* Mediator.

The second goal of my work was to determine whether *S. pombe* Ell1 and Eaf1 function as components of larger complexes, possibly related to SEC or LEC. In Chapter 3, I demonstrate that *S. pombe* Ell1 and Eaf1 associate with a new protein that is a 'sequence orphan', having no substantial similarity to any known protein. We refer to this protein as 'ELL binding protein 1 or Ebp1' as it interacts directly with Ell1. Although I was unable to detect Ebp1's effect on transcription elongation *in vitro*, results of my ChIP-seq experiments show that Ell1, Eaf1 and Ebp1 colocalize along with Pol II genomewide. It is consistent with the idea that they function together.

In Chapter 4, I use *ell1Δ*, *eaf1Δ* and *ebp1Δ* deletion mutant strains to study the consequences of loss of these proteins in cells, by studying their growth phenotypes in various conditions and by performing RNA-seq experiments and PRO-Seq experiments. I also performed a large scale genetic screen called 'Synthetic Genetic Array' (SGA) to identify genes that genetically interact with the genes encoding Ell1, Eaf1 and the Ebp1, as described in Chapter 5. Some preliminary data also suggests role of Ell1 in subtelomeric heterochromatin formation, which has been described in more detail in Chapters 4 and 5.

CHAPTER 2: Defining the *S. pombe* Mediator

2.1 Introduction

The *S. pombe* mediator has been shown to contain orthologs of all of the *S. cerevisiae* Head and Middle subunits except Med9; however, subunits corresponding to the Tail module (MED2, MED3, MED5, MED15 and MED 16) were not found to be associated with this complex in *S. pombe* (Beve et al., 2005b; Spahr et al., 2001). The Tail module is thought to be important for the interaction of mediator with DNA binding transcriptional activators (Myers et al., 1999). Even though *S. pombe* Mediator appears to lack the Tail domain, the genome of *S. pombe* encodes a putative MED15 ortholog (Linder et al., 2008b) that had not been shown to be associated with Mediator but was reported to exist in a protein complex along with Hrp1, a CHD1 family ATP dependent chromatin remodeling protein (Khorosjutina et al., 2010). Accordingly, it had been suggested that Mediator may influence gene chromatin structure by directly influencing the activity of the Hrp1 protein. Until recently, structural analysis of the *S. pombe* Mediator also only revealed the presence of a Head and Middle domain (Elmlund et al., 2006).

We wanted to know whether we could use *S. pombe* to study the functional interplay between Mediator and super elongation complex, as Ell1 and Eaf1 are also present in *S. pombe*. To address that question we first needed define the composition of *S. pombe* mediator. We were particularly interested in determining whether *S. pombe* mediator might be associated with a subunit that is related to the metazoan SEC-binding

subunit MED26, but so divergent in sequence as to be difficult to detect bioinformatically. I started by tagging few subunits (Med4 and Med 7) with endogenous Flag tag and performing Flag immunoprecipitations followed by MudPIT mass spectrometry. The Tail subunit Med15 was seen to be associated with the purified mediator complex. We also identified two 'sequence orphan' proteins that always copurified with the mediator complex. Analysis of their sequence showed a weak sequence similarity with Mediator subunits called Med 2 and Med 9. Notably Med 2 is a component of the Tail module. Though we identified neither a Med26-like subunit nor any SEC components in association with Mediator, we were able to provide the first comprehensive definition of the *S. pombe* Mediator. We collaborated with Francisco J. Asturias' Lab, which obtained a 4.4 Å cryo-electron microscopy map of the *S. pombe* Mediator in which conserved Mediator subunits are individually resolved (Tsai et al., 2017). The structural studies further helped to resolve the role of essential Med14 subunit as a central backbone that connects the Mediator Head, Middle and Tail modules.

2.2 Identification of Mediator subunits in *S. pombe*

To characterize the composition of Mediator complex in *S. pombe* I first selected two of its known subunits, Med7 and Med4, and endogenously tagged them with a 3X Flag tag at their C termini. After growing cells in rich media, Med7-FLAG and Med4-FLAG associated proteins were purified from lysates using anti-FLAG agarose immunoaffinity

chromatography. Mass spectrometry using multi-dimensional protein identification technology (MudPIT) was used to identify the proteins copurifying with Med4 and Med7.

Chromatography-based proteomic techniques like Multidimensional Protein Identification Technology (MudPIT) offer significant advantages over gel-based proteomic approaches for analysis of large, relatively fragile complexes such as the *S. pombe* Mediator. After digestion, protein samples are loaded directly onto a triphasic microcapillary column packed with reversed phase, strong cation exchange, and reversed phase HPLC grade materials. This column is then placed directly in-line with a tandem mass spectrometer that generates data to determine the total protein content of the original sample. This approach often allows detection of lower abundance proteins and/or proteins that are not readily visualized in SDS gels among the background of contaminant proteins that are often found in immunopurified samples (Wolters et al., 2001). MudPIT mass spectrometry was instrumental for defining the full complement of mammalian Mediator subunits (Sato et al., 2004).

Table 2.1 shows a list of all the Mediator subunits units identified from mass spectrometry. It is interesting to note that Med15 ortholog in *S. pombe*, that was earlier reported not to be associated with Mediator also copurifies in all the samples tested. Apart from the annotated subunits and Med15, there are two sequence orphans that also show up in all the samples in very high amounts. None of these subunits were present in the control samples that were FLAG-immunopurified proteins from the standard parental 972h- strain.

Also, various subunits of the RNA Polymerase II were pulled down along with the Mediator suggesting a portion of it was purified as a holoenzyme (data not shown).

Common Name	NCBI Locus Tag	MED7-FLAG_1 dNSAF	MED7-FLAG_2 dNSAF	MED7-FLAG_3 dNSAF	MED4-FLAG_1 dNSAF	MED4-FLAG_2 dNSAF	MED4-FLAG_3 dNSAF	FLAG-cont_1 dNSAF	FLAG-cont_2 dNSAF	FLAG-cont_3 dNSAF
MED1	SPAC2F7.04	0.0067	0.0067	0.0499	0.0015	0.0023	0.0207	X	X	X
sequence orphan (MED2)	SPCC4F11.03c	0.0039	0.0020	0.0145	0.0008	0.0018	0.0087	X	X	X
MED4	SPBC1105.06	0.0191	0.0165	0.0560	0.0035	0.0053	0.0277	X	X	X
MED6	SPAC1002.15c	0.0278	0.0540	0.0269	0.0058	0.0111	0.0136	X	X	X
MED7	SPBC14F5.08	0.0283	0.0182	0.0164	0.0048	0.0066	0.0113	X	X	X
MED8	SPBC21.04	0.0225	0.0386	0.0200	0.0120	0.0119	0.0101	X	X	X
sequence orphan (MED9)	SPAC24C9.04	0.0359	0.0560	0.0315	0.0140	0.0179	0.0233	X	X	X
MED10	SPBC31F10.09c	0.0065	0.0061	0.0061	0.0029	0.0037	0.0064	X	X	X
MED11	SPAC644.10	0.0299	0.0480	0.0357	0.0080	0.0125	0.0273	X	X	X
MED14	SPBC1A4.10c	0.0099	0.0187	0.0368	0.0046	0.0052	0.0200	X	X	X
MED15	SPBC146.01	0.0018	0.0015	0.0139	0.0002	0.0009	0.0091	X	X	X
MED17	SPBC31F10.04c	0.0154	0.0209	0.0460	0.0072	0.0051	0.0202	X	X	X
MED18	SPAC5D6.05	0.0231	0.0283	0.0179	0.0046	0.0074	0.0073	X	X	X
MED19	SPCC1450.05c	0.0161	0.0329	0.0189	0.0042	0.0038	0.0159	X	X	X
MED20	SPAC17G8.05	0.0104	0.0260	0.0257	0.0030	0.0034	0.0088	X	X	X
MED21	SPBC1604.10	0.0140	0.0076	0.0134	0.0031	0.0032	0.0082	X	X	X
MED22	SPAC29A4.07	0.0343	0.0871	0.0490	0.0120	0.0196	0.0191	X	X	X
MED27	SPAC17C9.05c	0.0040	0.0040	0.0277	0.0010	0.0024	0.0161	X	X	X
MED31	SPCP31B10.03c	0.0102	0.0130	0.0050	0.0010	0.0022	0.0058	X	X	X
MED12	SPAC688.08	0.0028	0.0024	0.0083	0.0005	0.0007	0.0048	X	X	X
MED13	SPAC589.02c	0.0036	0.0036	0.0125	0.0008	0.0013	0.0081	X	X	X
CDK8	SPAC23H4.17c	0.0058	0.0041	0.0048	0.0010	0.0013	0.0038	X	X	X
Cyclin C	SPBC12D12.06	0.0014	0.0004	0.0024	0.0003	0.0004	0.0010	X	X	X

Table 2.1 Identification of *S. pombe* Mediator subunits in Med4-FLAG and Med7-FLAG SpMED preparations by MudPIT mass spectrometry.

This table shows the distributed normalized spectral abundance factor (dNSAF) for each Mediator subunit copurifying with *S. pombe* FLAG- Med4, and Med7. Control samples are FLAG-immunopurified proteins from the standard parental 972h- strain. dNSAF values provide a measure of the relative amount of the same protein across several different samples, but only a rough estimate of the relative amounts of each protein detected in a MudPIT data set.

2.3 Association of Med15 with Mediator

Med15 protein had been previously suggested not to be a component of the *S. pombe* Mediator; however, our Flag purifications using FLAG- Med4 and Med7 identified it as a likely Mediator-associated protein.

To further confirm this interaction, I Flag tagged endogenous Med15 and performed Flag purification followed by MudPIT analysis. Table 2.2 (lanes 1-2) shows that most of the annotated subunits of Mediator along with the two sequence orphans corresponding to MED2 and MED9 copurified with Med15. The yield of Mediator subunits was relatively low (based on dNSAF) in MED15-purified fractions, an observation that could be explained if MED15 tends to dissociate from Mediator or if the C-terminal tag is relatively inaccessible when MED15 is assembled into Mediator. This could provide an explanation why Mediator subunits were not detected in previous analyses of MED15-associated proteins using methods less sensitive than MudPIT.

It is interesting to note that the chromatin remodeling protein, Hrp1, which was previously reported to form a complex with Med15, was also identified in these samples (Table 2.2, last row). However, the protein has almost similar abundance in the control samples having no Flag tag (last row, lanes 4-6), raising the possibility that Hrp1 might bind non-specifically to FLAG-agarose rather than being a bona fide Med15- or Mediator-binding protein.

Common Name	NCBI Locus Tag	MED15-FLAG_1 dNSAF	MED15-FLAG_2 dNSAF	MED2-FLAG_1 dNSAF	FLAG-cont_1 dNSAF	FLAG-cont_2 dNSAF	FLAG-cont_3 dNSAF
MED1	SPAC2F7.04	0.0001	0.0000	0.0250	X	X	X
sequence orphan (MED2)	SPCC4F11.03c	0.0007	0.0001	0.0303	X	X	X
MED4	SPBC1105.06	0.0003	0.0004	0.0614	X	X	X
MED6	SPAC1002.15c	0.0000	0.0002	0.0248	X	X	X
MED7	SPBC14F5.08	0.0001	0.0002	0.0279	X	X	X
MED8	SPBC21.04	0.0004	0.0004	0.0185	X	X	X
sequence orphan (MED9)	SPAC24C9.04	0.0002	0.0006	0.0943	X	X	X
MED10	SPBC31F10.09c	0.0008	0.0002	0.0215	X	X	X
MED11	SPAC644.10	0.0005	0.0005	0.0602	X	X	X
MED14	SPBC1A4.10c	0.0002	0.0000	0.0323	X	X	X
MED15	SPBC146.01	0.0209	0.0072	0.0303	X	X	X
MED17	SPBC31F10.04c	0.0002	0.0001	0.0388	X	X	X
MED18	SPAC5D6.05	X	0.0001	0.0163	X	X	X
MED19	SPCC1450.05c	0.0003	0.0004	0.0202	X	X	X
MED20	SPAC17G8.05	0.0001	0.0002	0.0199	X	X	X
MED21	SPBC1604.10	0.0005	0.0002	0.0256	X	X	X
MED22	SPAC29A4.07	0.0003	0.0003	0.0410	X	X	X
MED27	SPAC17C9.05c	0.0007	0.0002	0.0277	X	X	X
MED31	SPCP31B10.03c	0.0001	0.0001	0.0188	X	X	X
MED12	SPAC688.08	0.0000	0.0000	0.0043	X	X	X
MED13	SPAC589.02c	0.0002	X	0.0087	X	X	X
CDK8	SPAC23H4.17c	X	X	0.0023	X	X	X
Cyclin C	SPBC12D12.06	X	X	0.0008	X	X	X
hrp1	SPAC1783.05	0.0261	0.0218	0.0128	0.0245	0.078	0.0525

Table 2.2 Identification of *S. pombe* Mediator subunits in Med15-FLAG and Med2-FLAG SpMED preparations by MudPIT mass spectrometry.

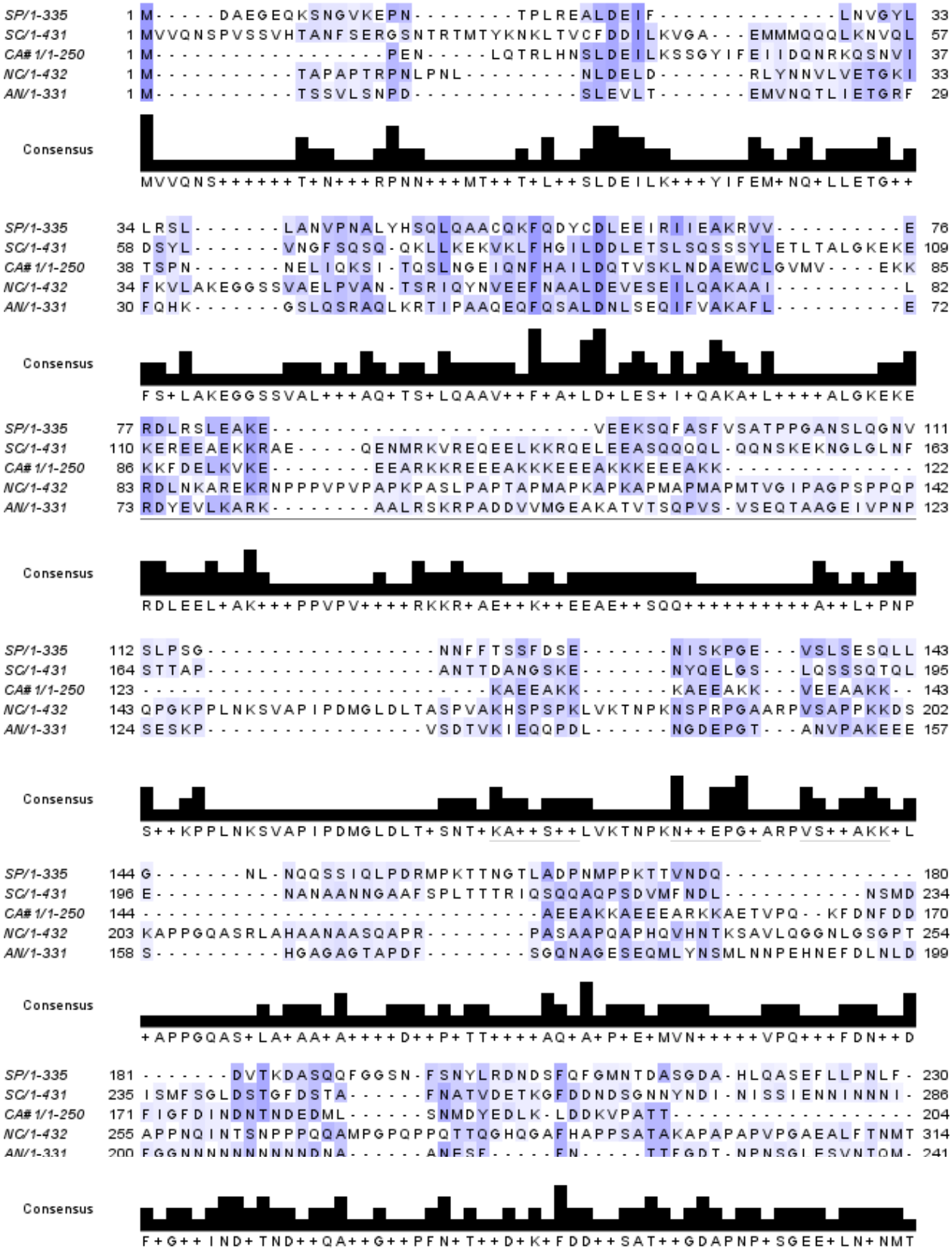
Similar to Table 2.1, this table shows the dNSAF for each Mediator subunit copurifying with *S. pombe* FLAG-Med15, and putative Med2. Control samples are FLAG-immunopurified proteins from the standard parental 972h- strain.

2.4 Characterization of 'sequence orphan' Mediator subunits

Two sequence orphans copurify in all the Mediator preparations, that had never been shown experimentally to be Mediator subunits. To confirm the association of these

putative Mediator subunit, I flag-tagged one of them (endogenous *SpSPCC4F11.03c*) at C terminal and did Flag purification followed by MudPIT analysis. And it pulls out all known Mediator subunits with Med15 and the other sequence orphan (Table 2.2). Taken together, that suggests all are Mediator subunits.

We could not use standard psi-blast to identify sequence orphans as orthologs of Mediator subunits in other species. Using sophisticated multi-sequence alignment approaches, Bourbon had proposed that the two sequence orphans might be Med2 and Med9, but relationship was quite distant, and it wasn't at all clear whether they really were (Bourbon, 2008). Figure 2.1 and 2.2 show a simplified version of the alignment published by Bourbon of Med2 and Med9 across many species, also having putative Med2 and Med9 of *S. pombe* (encoded by *SPCC4F11.03c* and *SPAC24C9.04* in *S. pombe*, respectively).


SP/1-335 1 M-----DAEGEQKSNGVKEPN--------TPLREALDEIF--------------LNVGYL 33
SC/1-431 1 MVVQNSPVSSVHTANFSERGSNTRTMTYKNKLTVCFDDILKVGA---EMMMQQQLKNVQL 57
CA#1/1-250 1 M----------------PEN-----LQTRLHNSLDEILKSSGYIFEIIDQNRKQSNVI 37
NC/1-432 1 M----------TAPAPTRPNLPNL---------NLDELD-------RLYNNVLVETGKI 33
AN/1-331 1 M----------TSSVLSNPD-------------SLEVLT-------EMVNQTLIETGRF 29
Consensus
MVVQNS++++++T+N+++RPNN+++MT++T+L++SLDEILK+++YIFEM+NQ+LLETG++
SP/1-335 34 LRSL-------LANVPNALYHSQLQAACQKFQDYCDLEEIRIIEAKRVV----------E 76
SC/1-431 58 DSYL-------VNGFSQSQ-QKLLKEKVKLFHGILDDLETSLSQSSSYLETLTALGKEKE 109
CA#1/1-250 38 TSPN-------NELIQKSI-TQSLNGEIQNFHAILDQTVSKLNDAEWCLGVMV----EKK 85
NC/1-432 34 FKVLAKEGGSSVAELPVAN-TSRIQYNVEEFNAALDEVESEILQAKAAI----------L 82
AN/1-331 30 FQHK-------GSLQSRAQLKRTIPAAQEQFQSALDNLSEQIFVAKAFL----------E 72
Consensus
FS+LAKEGGSSVAL+++AQ+TS+LQAAV++F+A+LD+LES+I+QAKA+L++++ALGKEKE
SP/1-335 77 RDLRSLEAKE----------------------------VEEKSQFASFVSATPPGANSLQGNV 111
SC/1-431 110 KEREEAEKKRAE-----QENMRKVREQEELKKRQELEEASQQQQL-QQNSKEKNGLGLNF 163
CA#1/1-250 86 KKFDELKVKE--------EEARKKREEEAKKKEEEAKKKEEEAKK-------------- 122
NC/1-432 83 RDLNKAREKRNPPPVPVPAPKPASLPAPTAPMAPKAPKAPMAPMAPMTVGIPAGPSPPQP 142
AN/1-331 73 RDYEVLKARK--------AALRSKRPADDVVMGEAKATVTSQPVS-VSEQTAAGEIVPNP 123
Consensus
RDLEEL+AK+++PPVPV++++RKKR+AE++K++EEAE++SQQ++++++++++A++L+PNP
SP/1-335 112 SLPSG----------------NNFFTSSFDSE-------NISKPGE---VSLSESQLL 143
SC/1-431 164 STTAP----------------ANTTDANGSKE-------NYQELGS---LQSSSQTQL 195
CA#1/1-250 123 ----------------------KAEEAKK-------KAEEAKK---VEEAAKK-- 143
NC/1-432 143 QPGKPPLNKSVAPIPDMGLDLTASPVAKHSPSPKLVKTNPKNSPRPGAARPVSAPPKKDS 202
AN/1-331 124 SESKP----------------VSDTVKIEQQPDL------NGDEPGT---ANVPAKEEE 157
Consensus
S++KPPLNKSVAPIPDMGLDLT+SNT+KA++S++LVKTNPKN++EPG+ARPVS++AKK+L
SP/1-335 144 G-------NL-NQQSSIQLPDRMPKTTNGTLADPNMPPKTTVNDQ---------------- 180
SC/1-431 196 E----------NANAANNGAAFSPLTTTRIQSQQAQPSDVMFNDL----------NSMD 234
CA#1/1-250 144 ----------------------------AEEAKKAEEEARKKAETVPQ--KFDNFDD 170
NC/1-432 203 KAPPGQASRLAHAANAASQAPR--------PASAAPQAPHQVHNTKSAVLQGGNLGSGPT 254
AN/1-331 158 S----------HGAGAGTAPDF--------SGQNAGESEQMLYNSMLNNPEHNEFDLNLD 199
Consensus
+APPGQAS+LA+AA+A++++D++P+TT++++AQ+A+P+E+MVN+++++VPQ+++FDN++D
SP/1-335 181 -------DVTKDASQQFGGSN-FSNYLRDNDSFQFGMNTDASGDA-HLQASEFLLPNLF- 230
SC/1-431 235 ISMFSGLDSTGFDSTA------FNATVDETKGFDDNDSGNNYNDI-NISSIENNINNNI- 286
CA#1/1-250 171 FIGFDINDNTNDEDML------SNMDYEDLK-LDDKVPATT---------------- 204
NC/1-432 255 APPNQINTSNPPPQQAMPGPQPPQTTQGHQGAFHAPPSATAKAPAPAPVPGAEALFTNMT 314
AN/1-331 200 FGGNNNNNNNNNNDNA------ANFSF-----FN-----TTFGDT-NPNSGLFSVNTQM- 241
Consensus
F+G++IND+TND++QA++G++PFN+T++D+K+FDD++SAT++GDAPNP+SGEE+LN+NMT

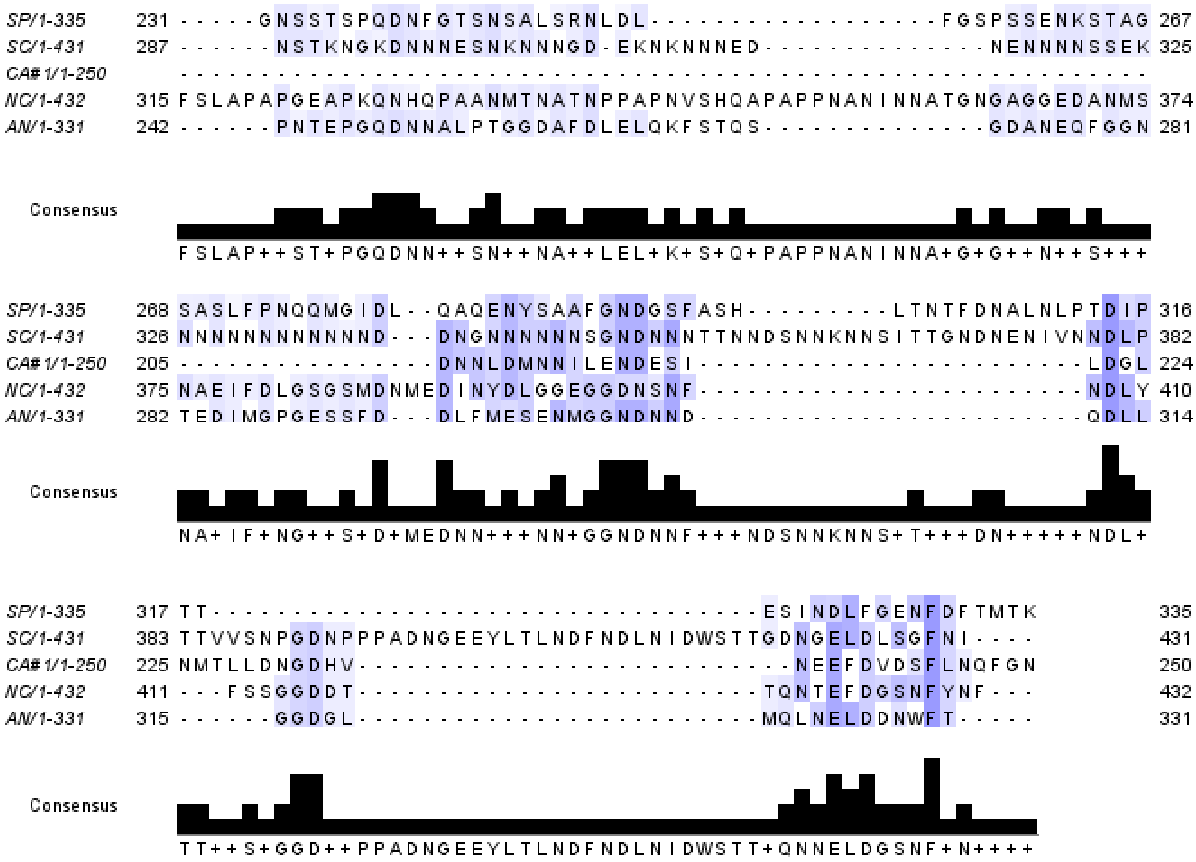


Figure 2.1 Structural conservation of predicted Mediator subunit Med2 among yeast.

Figure derived from (Bourbon, 2008; Tsai et al., 2014). Primary sequences of few fungal (*Schizosaccharomyces pombe*(SP), *Saccharomyces cerevisiae*(SC), *Candida albicans*(CA), *Neurospora crassa*(NC), *Aspergillus nidulans*(AN)) MED2 subunits identified through PSIBlast or TBlastN analyses were aligned using MAFFT and colored with JALVIEW based on conservation across these species. Bottom panel shows the consensus sequence predicted based on conservation

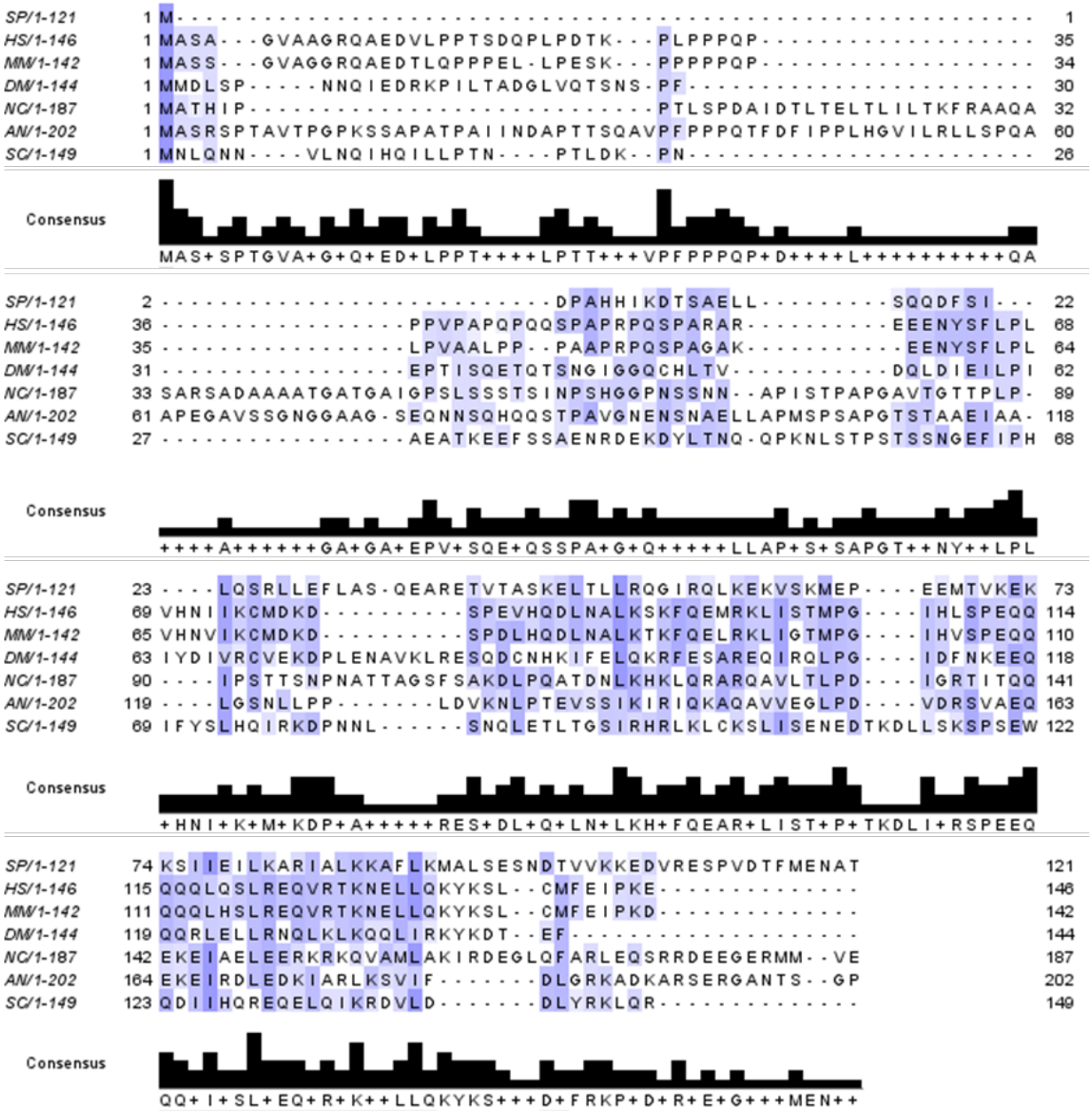


Figure 2.2 Structural conservation of predicted Mediator subunit Med9.

Figure derived from (Bourbon, 2008; Tsai et al., 2014). Primary sequences of few fungal (*Schizosaccharomyces pombe*(SP), *Saccharomyces cerevisiae*(SC), *Candida albicans*(CA), *Neurospora crassa*(NC), *Aspergillus nidulans*(AN)) MED2 subunits identified through PSIBlast or TBlastN analyses were aligned using MAFFT and colored with JALVIEW based on conservation across these species. Bottom panel shows the consensus sequence predicted based on conservation

2.5 Structural studies on *S. pombe* Mediator

By defining for the first time the complete repertoire of *S. pombe* Mediator subunits, our proteomic analysis of *S. pombe* Mediator provided information essential for interpretation of cryo-electron microscopy-based structural studies *S. pombe* Mediator performed in the laboratory of Francisco Asturias. Cryo-EM experiments performed by Kuang-Lei Tsai in Asturias Lab (Department of Integrative Structural and Computational Biology, The Scripps Research Institute). These cryo-EM studies localized specific Mediator subunits and defined in unprecedented detail the molecular architecture of the Mediator complex. As discussed above, my work led to identification of the ORFs corresponding to Med2, Med 9 and Med 15 as *S. pombe* Mediator subunits. The high-resolution structure of Middle module showed that Med9-like molecule was present, supporting the ID of one the sequence orphan discussed above as Med9In addition, a possible ortholog of human Mediator MED27 was also identified. It is generally acknowledged that SpMED does not include Tail module subunits corresponding to Med5 and Med16. To determine whether the *S. pombe* putative Med2 subunit was, as predicted by sequence homology, a component of a previously unrecognized *S. pombe* Tail module, I generated a generated a *med2Δ* strain in a Med7-TAP, *med13Δ* genetic background that allowed purification of *S. pombe* Mediator suitable for EM. Mediator purified from this strain was used in difference mapping experiments, in which one compares the structures of the complete Mediator complex to Mediator lacking specific subunit (in this case Med2). EM analysis of *med2Δ* SpMED revealed a loss of partially ordered Tail density in a position corresponding to the Med2 position in *S. cerevisiae*

(Figure 2.3A), arguing that the location of this subunit is conserved across species. We were unable to obtain a *med15Δ* strain; however, Med15 in *S. cerevisiae* interacts directly with Med2 (Beve et al., 2005), and diffuse density in 2D and 3D SpMED maps is consistent with the existence of a poorly ordered Med2/Med15 Tail.

Med27 is found in metazoan Mediator and, as I have shown, in *S. pombe* Mediator. Because there are no known Med27 homologs in *S. cerevisiae*, its location had not been defined in previous structural studies of budding yeast Mediator. Accordingly, it was of particular interest to exploit *S. pombe* to define the location of Med27. I therefore generated a *med27Δ* strain that could be suitable for purification of Mediator lacking Med27. Difference mapping experiments performed with *med27Δ* SpMED particles localized Med27 density to the distal end of Med20, connecting the Med18-Med20 Head jaw to Med17 (Figure 2.3B). These results are consistent with biochemical studies showing that human Med27 can interact with several Head module subunits, including Med17, Med18, Med20 (Tsai et al., 2014). An extension of Med27 to the Tail was also apparent in the SpMED class averages (Figure 2.3C), suggesting it could from a weak connection between purported Med27 density in the Head and Tail modules, but further experiments will be needed to confirm this model.

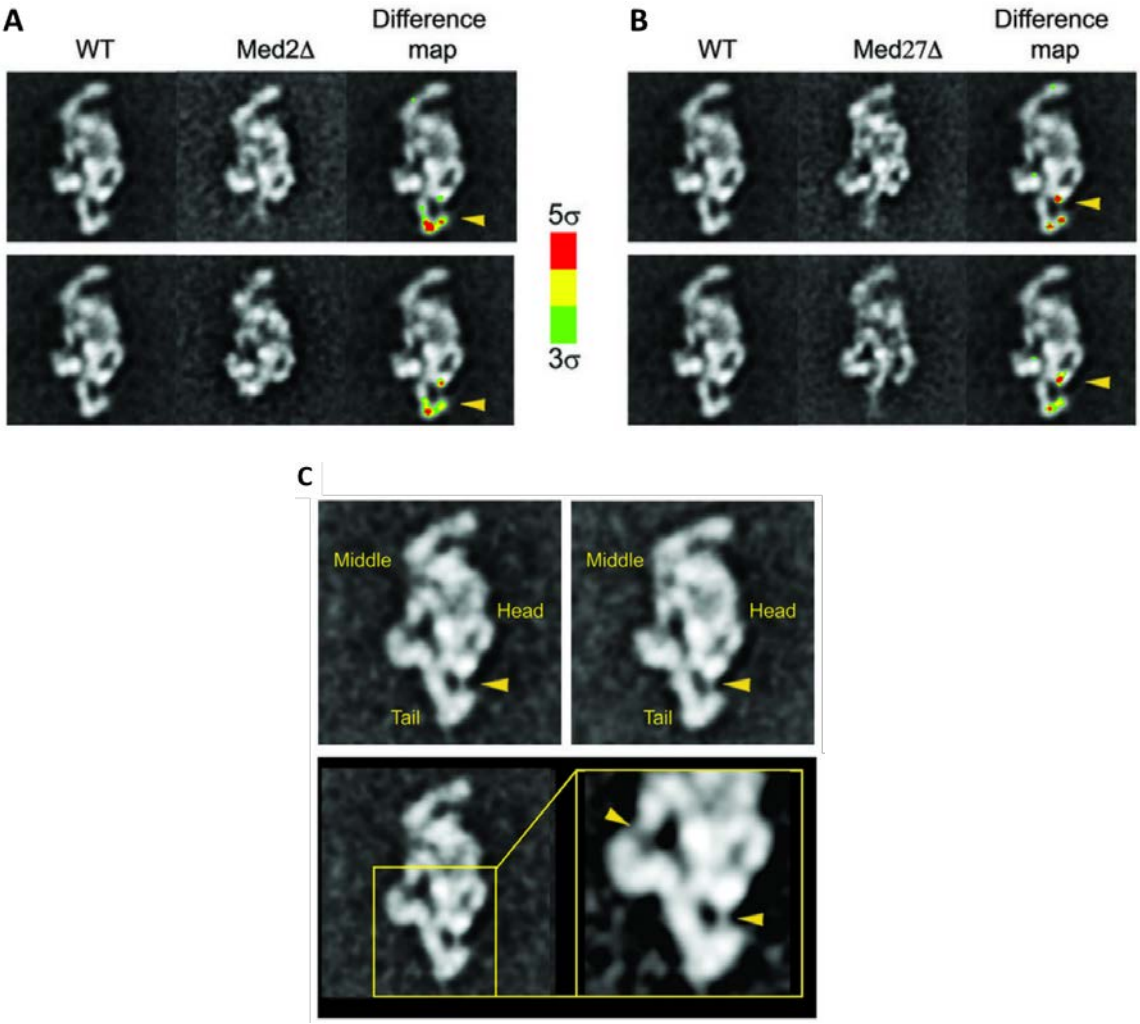


Figure 2.3 Subunit localization of *S. pombe* Mediator.

(**A, B**), 2D class averages for wild-type, *med2Δ* (A) and *med27Δ* (B) SpMED, and color-coded (by standard deviation values from the average) difference maps indicating the position of Med2 and Med27 (highlighted by yellow arrowheads). **(C)**, Wild-type SpMED 2D class averages and a close-up showing a Med27–Tail connection (bottom-right arrowhead) comparable to the connection between the Med4–Med9 four-helix bundle and the rest of the Middle module (top-left arrowhead). (reprinted with permission from Francisco Asturias)

2.6 Discussion

Studies done to investigate the mechanism of action of Mediator from yeast to higher organisms revealed that Mediator can directly interact with both the Pol II initiation machinery and the DNA binding transcription factors. Mediator also helps in transcription elongation by (1) controlling phosphorylation of the heptad repeats of the

Pol II CTD, (2) by recruiting Pol II elongation factors and other pre-mRNA processing factors and (3) by overcoming the activities of the negatively regulating factors.

Some Mediator subunits can be deleted without loss of viability. Although not all subunits are essential for growth in rich media, many show phenotypes under specific growth conditions. For example, the Med15 gene was identified as GAL11, which is required for optimal growth when galactose is carbon source and for induction of galactose-inducible genes (Nogi and Fukasawa, 1980). In *S. cerevisiae*, the Mediator Tail module interacts with many gene-specific transcription factors (e.g. Gal4 and Gcn4). Except for Med15, the *S. pombe* genome was not known to encode any obvious homologs to budding yeast Tail subunits. Notably, the *S. pombe* homolog of Med15 had been reported to be associated with the chromodomain-containing chromatin remodeling enzyme Hrp1 complex and was not known to associate with the SpMED.

Our Mediator purifications and MudPIT mass spectrometry analyses, along with the EM structures from Asturias and colleagues, show that the fission yeast Mediator does indeed have an attenuated Tail module. Homologs of the Tail subunits Med2 and Med15 were identified in MudPIT analyses of the purified SpMediator. EM analyses reveal densities in SpMediator structure corresponding to a Med2/Med15 Tail. Of note, initial analysis also suggests that Med27, which has been assigned to the Head module in metazoan Mediator (Tsai et al., 2014) appears to connect the jaw region (Med18/Med20) to Tail.

Subunits of the *S. cerevisiae* Tail module have been shown to be substoichiometric relative to other core Mediator components in purified Mediator preparations (Myers et al., 1998), and it is therefore possible that the Tail module is

found in only a subset of budding yeast Mediator complexes. Others have also demonstrated that the Tail module of budding yeast Mediator can function as a separate entity in certain mutant backgrounds (Zhang et al., 2004). It is hence possible that *S. pombe* Tail module also flexibly associates with other Mediator subunits. It is worth noting that although *S. pombe* Med15 clearly binds Mediator, other Mediator subunits are present at rather low levels in FLAG-immunopurified fractions from Med15-FLAG-expressing *S. pombe*. This observation raises the possibilities that a substantial fraction of Med15 is present in free form, unbound to the Mediator, or is associated with some other protein(s) / complex(es). It is possible that the Tail module associates with the Mediator complex only in the presence of Activator proteins. Also, components of the Kinase module seem not to be associated with the Mediator pulled down by Med15 immunoprecipitation, suggesting that association of the Tail module and the Kinase module with Mediator could be mutually exclusive in *S. pombe*. We did not check the Mediator composition under different growth conditions or cell cycle stages, but there is a possibility that the subunits' association change under different conditions. We did not find associated proteins that could function as a putative Med26 in *S. pombe*, nor did we detect any association between the Mediator complex and elongation factors like Ell1, Eaf1 or Cdk9. Thus, other experimental strategies will be required to address the question whether and how Mediator plays a role in regulation of Ell1, Eaf1, and Cdk9 function in *S. pombe*.

Cryo-EM studies by Francisco Asturias' group on Mediator and holoenzyme reveal crucial details, at a near-atomic resolution. The 4.3 Å cryo-EM structure of *S. pombe* Mediator provides first near-atomic view of Mediator in its native conformation. Well-ordered Head and Middle modules, and a disordered Tail are organized around an

extended, central Med14 that facilitates all inter-module contacts. Changes in Med14 structure facilitates large-scale changes in Mediator conformation and makes possible polymerase contacts and enhanced CTD interaction. Matching of the polymerase portion of the holoenzyme and preinitiation complex cryo-EM structures indicates that Mediator rearrangements would be essential to stabilize the preinitiation complex and facilitate targeting of the CTD for phosphorylation by TFIIK. It is speculated that factors involved in regulation of activated transcription might act by exploiting the Mediator conformational changes, for example, through internal flexibility of the Head or through Tail rearrangements facilitated by the Med14 C-terminal domain.

CHAPTER 3: Identification of a new ELL-EAF interacting partner

3.1 Introduction

The *ELL* gene was first identified as an *MLL* translocation partner in acute myeloid leukemia (Thirman et al., 1994). Subsequently, the Ell protein was purified from rat liver extract as an activity that can stimulate the rate of RNA Polymerase II transcription elongation *in vitro* (Shilatifard et al., 1996). ELL can bind directly to elongating Pol II and hence can help to promote productive elongation (Shilatifard et al., 1997c). Mammalian ELL binds directly to EAF1 and EAF2 (ELL associated factor), which act as positive regulators of ELL transcription *in vitro* (Kong et al., 2005).

ELL and EAF in higher eukaryotes can also be present as part of larger complexes like the Super elongation complex (SEC). The SEC consists of the Pol II elongation factors ELL and P-TEFb and several frequent MLL translocation partners including the AF4 family members AF4 and AFF4, and the AF9 family members AF9 and ENL.

Most of the components of the Pol II transcription machinery are conserved from yeast to humans, but early attempts to find orthologs of ELL and EAF in yeast were unsuccessful, suggesting that ELL-EAF function might have evolved only in higher eukaryotes having larger genes and genomes. More recent studies from our laboratory identified a *Schizosaccharomyces pombe* RNA polymerase II elongation factor with similarity to the metazoan transcription factor ELL (Banks et al., 2007). Using bioinformatic tools, genes encoding proteins with weak similarity to ELL and EAF were identified in the *S. pombe* genome. Recombinant proteins encoded by these predicted ORFs, *SPBP23A10.14c* and *SPCC1223.10c*, were shown to be capable of stimulating

elongation by *S. pombe* Pol II and accordingly were annotated *ell1* and *eaf1*, respectively. No similar proteins could be identified in *S. cerevisiae* and related fungi. Notably, *S. cerevisiae* is different from other fungi and from higher eukaryotes as it lacks several enzyme systems, including those needed for regulating many splicing processes and for RNAi.

Since *S. pombe* has ELL, EAF and P-TEFb orthologs, we considered the possibility that it might form larger elongation complexes such as the SEC. We searched the *S. pombe* genome and couldn't find anything that looks obviously like an AF4 or AF9 family member but the Ell1/Eaf1-associated proteins maybe distantly related to AF4 or AF9 proteins and hence biochemically we may have a better chance of finding them. I tagged Ell1 and Eaf1 with Flag tags and did Immunoprecipitation followed by MudPIT. We identified a previously uncharacterized protein encoded by the gene *SPAC6G9.15c* to be associated with the both Ell1 and Eaf1. Biochemical experiments with recombinant proteins showed that the protein encoded by *SPAC6G9.15c* (Ebp1) can bind directly to Ell1 and form a stable complex with Ell1 and Eaf1. Chip seq experiments showed that Ell1, Eaf1 and Ebp1 are co-recruited *in vivo* to genes having high Pol II occupancy. Ell1, Eaf1 and Ebp1 occupancy also correlate with P-TEFb (Cdk9) occupancy. Ell1 and Eaf1 together can stimulate elongation *in vitro* but the presence of Ebp1 had no noticeable effect on transcription under the conditions tested.

3.2 Ell1 and Eaf1 copurify with an uncharacterized gene product

To determine the proteins that can interact with Ell1 and Eaf1, I individually tagged the endogenous protein with a C-terminal 3X FLAG tag. The cells were grown in rich media and Ell1-FLAG and Eaf1-FLAG associated proteins were purified using anti-FLAG agarose immunoaffinity chromatography. Mass spectrometry using multi-dimensional protein identification technology (MudPIT) was used to identify the proteins copurifying with Ell1 and Eaf1 . As shown in Table 3.1 and 3.2, both Ell1 and Eaf1 copurifies with a previously uncharacterized protein encoded by the gene *SPAC6G9.15c*. Gene product of *SPAC6G9.15c* was annotated as a sequence orphan as it had no sequence homology to any other protein in another species. Because it interacts directly with Ell1, we annotated this protein as "ELL binding protein 1" or "Ebp1".

	Ell1-FLAG									No Tag control								
	1			2			3			1			2			3		
	P	S	dNSAF	P	S	dNSAF	P	S	dNSAF	P	S	dNSAF	P	S	dNSAF	P	S	dNSAF
Ell1	42	294	0.0066	38	344	0.0060	54	445	0.0429	0	0	X	0	0	X	0	0	X
Eaf1	13	44	0.0022	21	240	0.0090	19	857	0.0472	2	3	0.0001	0	0	X	0	0	X
SPAC6G9.15c	24	101	0.0024	31	145	0.0027	30	336	0.0180	0	0	X	0	0	X	0	0	X

Table 3.1 Ell1 copurifies with Eaf1 and *SPAC6G9.15c* gene product.

Cell lysates from *S. pombe* expressing Ell1-FLAG or untagged strain were purified by anti-FLAG immunochromatography and subjected to MudPIT mass spectrometry. The table shows the number of peptides (P) and spectra (S) for Ell1, Eaf1 and *SPAC6G9.15c* gene product detected in samples purified from cells expressing either Ell1-FLAG or untagged strain. *NSAF*, normalized spectral abundance factor. The number of spectra for a given protein detected in a MudPIT run has been shown to be a function of the protein's size and abundance (122); hence, comparison of the spectral counts provides a rough estimation of the relative abundance of different proteins across samples. NSAF is the spectral count for a given protein, normalized to the protein's length and the total number of spectra detected in the MudPIT run. NSAF is calculated using the equation

$$(NSAF)_k = \frac{(SpC / Length)_k}{\sum_{i=1}^{N} (SpC / Length)_i}$$

where *SpC* is the number of spectra detected, *k* denotes a specific protein, and *N* is all proteins detected.

	Eaf1-FLAG									No Tag control								
	1			2			3			1			2			3		
	P	S	dNSAF	P	S	dNSAF	P	S	dNSAF	P	S	dNSAF	P	S	dNSAF	P	S	dNSAF
Ell1	16	307	0.0131	32	392	0.0064	35	620	0.0395	0	0	X	0	0	X	0	0	X
Eaf1	20	319	0.0290	21	421	0.0147	24	507	0.0686	2	3	0.0001	0	0	X	0	0	X
SPAC6G9.15c	1	3	0.0001	20	68	0.0012	14	125	0.0085	0	0	X	0	0	X	0	0	X

Table 3.2 Eaf1 copurifies with Ell1 and *SPAC6G9.15c* gene product.

Similar to Table 3.1, Eaf1-FLAG or untagged strain were used to detect proteins associated with Eaf1 using anti-FLAG IP and MudPIT.

3.3 Ebp1 binds directly to Ell1 and forms a stable "ELL complex" with Ell1/Eaf1

To determine whether the interaction between Ell1, Eaf1, and Ebp1 is direct or not, I subcloned the Ell1, Eaf1 and SPAC6G9.15c ORFs into baculovirus vectors and expressed them in SF9 cells in several epitopes tagged forms. I expressed FLAG-Ell1, Myc-Eaf1 and HA-Ebp1 in various combinations as indicated in the figures, and performed anti anti-FLAG, anti-Myc, or anti-HA agarose chromatography. As shown in the anti-FLAG-Ell1 pull-downs in Fig. 3.1, FLAG-Ell1 could pull down both Eaf1 and Ebp1 when all three proteins were expressed together and when it was expressed with either Eaf1 or Ebp1 independently.

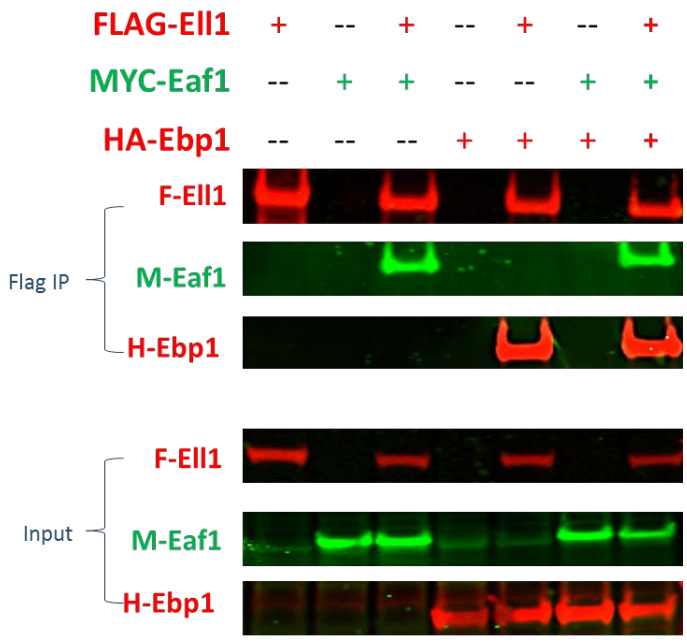


Figure 3.1 Ell1 interacts with both Eaf1 and Ebp1.

Sf9 cells coinfected with baculoviruses encoding FLAG -Ell1, Myc-Eaf1 and HA-Ebp1 in the combinations indicated in the figure were prepared and immunoprecipitations were carried out with the anti-FLAG agarose beads. Bound proteins were eluted with 250ng/µl FLAG peptide, analyzed by SDS-PAGE, and detected by either Western blotting using the specified antibody.

On the other hand, the interaction between Eaf1 and Ebp1 seems to be dependent on the presence of Ell1. Using the same lysates as were used in Figure 3.2, Myc or HA IPs were performed. As shown in Fig. 3.2, an interaction between Myc-Eaf1 and HA-Ebp1 was detected only in cells that also expressed Ell1.

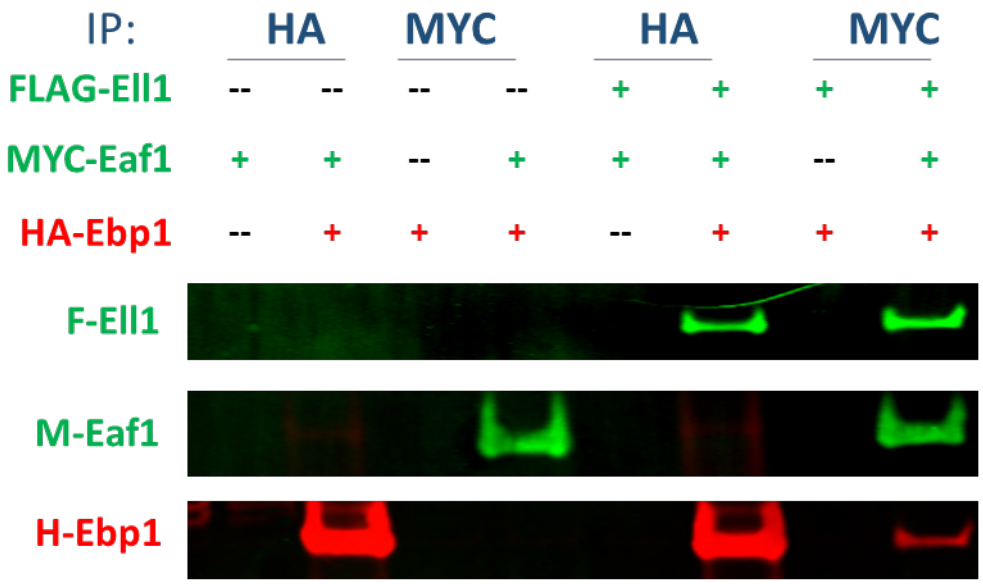


Figure 3.2 Ell1 is required for Eaf1 and Ebp1 to interact.

Similar to Figure 3.1, cells coinfected with baculoviruses encoding FLAG -Ell1, Myc-Eaf1 and HA-Ebp1 in the combinations indicated in the figure were prepared and immunoprecipitations were carried out with the indicated antibody. Bound proteins were eluted with HA or Myc peptide, analyzed by Western blotting using the specified antibody.

3.4 Ell1, Eaf1 and Ebp1 are co-recruited to genes *in vivo*

I hypothesized that since these proteins can bind to form a complex in vitro, they might also colocalize in the genome. We purified FLAG-Ell1, FLAG-Eaf1 and FLAG-Ebp1 from SF9 insect cells and used the recombinant proteins to prepare polyclonal antibodies in rabbits (Bio-Synthesis®). The specificity and appropriate antibody dilutions were confirmed using western blots (data not shown).

These polyclonal antibodies were then used to perform chromatin immunoprecipitations (ChIPs) using a wild type strain (Pem2) of *S. pombe*. Immunoprecipitated DNA was then analyzed by Next-Gen sequencing. To control for the specificity of ChIP signals, we performed parallel experiments in strains lacking genes encoding either Ell1, Eaf1 or Ebp1.

Figure 3.3A shows an IGV screenshot of 180kb region in *S. pombe* showing the peaks obtained from ChIP seq using α-Ell1, α-Eaf1 or α-Ebp1 antibodies in wildtype and corresponding deletion strains. We can see that there are various loci where Ell1, Eaf1 and Ebp1 give strong enrichment in the wildtype strain. These loci seem to overlap across for all three proteins, suggesting that Ell1, Eaf1 and Ebp1 co-localize in the genome. We see the enrichment of the peaks is drastically decreased in the deletion strains lacking the protein against which the antibody was raised. The majority of peaks overlapped annotated transcripts.

Data from replicates were highly correlated and were merged using samtools merge. Peaks were called using MACS2, using as background (input) the datasets generated by anti-Ell1, Eaf1, or Ebp1 ChIP of chromatin from *ell1Δ*, *eaf1Δ*, or *ebp1Δ* strains, respectively. Peak lists were filtered as follows. We removed the annotated antisense gene of the protein coding genes from the gene list. Genes that were mitochondrial in nature were also filtered out.

Figure 3.3B is a Venn diagram showing overlap between Ell1-, Eaf1-, and Ebp1 -occupied genes in wild type *S. pombe* (764 genes). Also, there is a subset of 670 genes that have colocalization of only Ell1 and Eaf1.

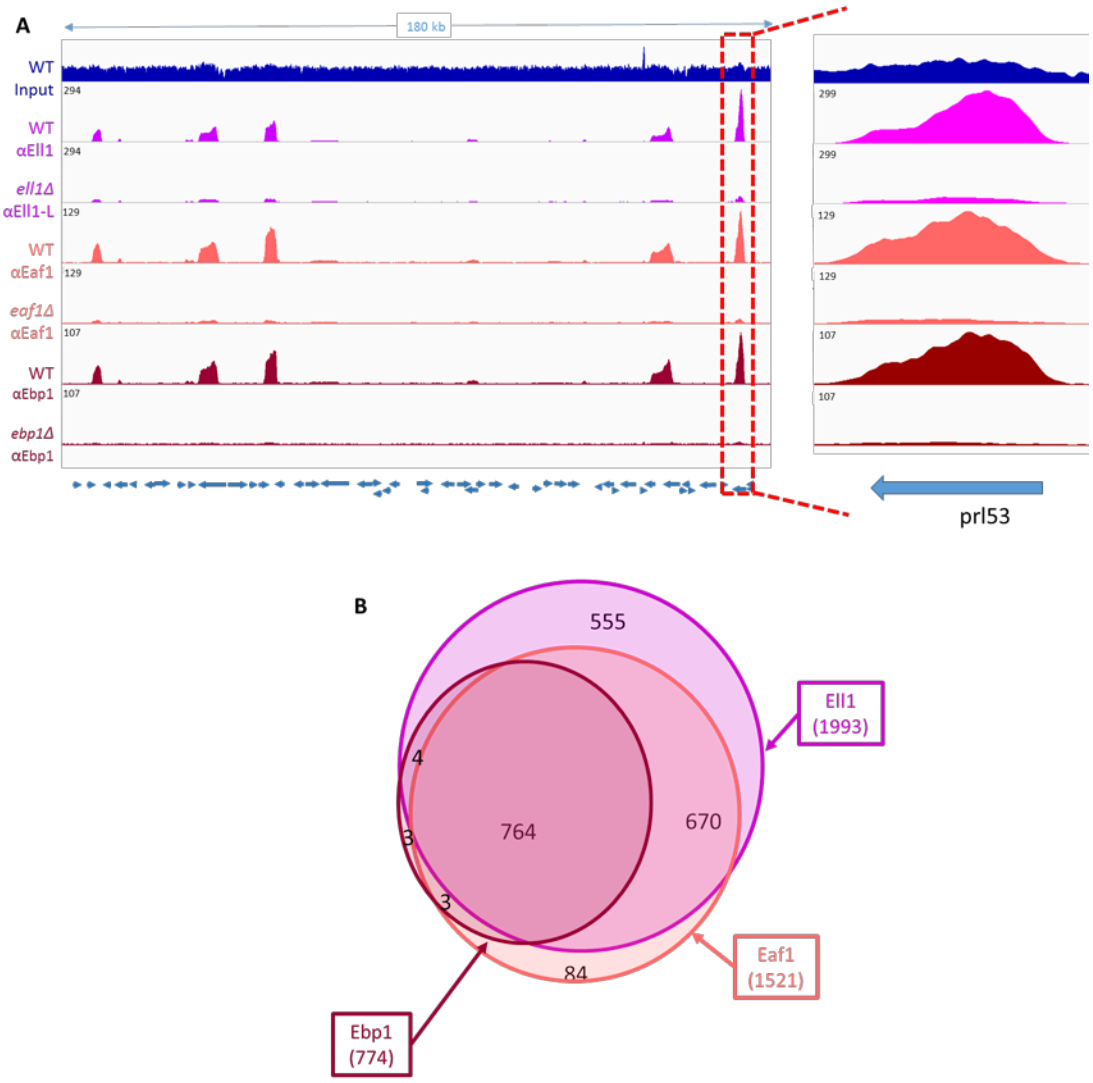


Figure 3.3 Ell1, Eaf1 and Ebp1 are co-recruited to genes *in vivo*.

A) IGV screenshot of 180kb region in *S. pombe*. ChIP-seq tracks corresponding to Ell1 (purple), Eaf1 (orange) and Ebp1 (brown) in Wildtype(WT) and corresponding deletion mutants are shown. Zooming into a peak of *prl53* gene shows that the enrichment of Ell1, Eaf1 and Ebp1 proteins on the gene. The arrows at the bottom represent gene directions and the input track is shown in blue. (B) Venn diagram showing overlap between Ell1-, Eaf1-, and Ebp1-occupied genes in wild type *S. pombe*. ChIP peaks corresponding to anti-Ell1, Eaf1 and Ebp1 were assigned to the closest gene and the three colors represent the number of genes with significant enrichment (50>FC>5, q-value<0.01)

3.5 Genes enriched by Ell1, Eaf1 and Ebp1 also have high Pol II

Ell1-EAF1 binds to Pol II in vitro and stimulates elongation. Accordingly, we expected that the ELL complex might co-localize with actively transcribing Pol II, so we asked whether Ell1, Eaf1, and Ebp1 occupancy is correlated with that of Pol II. To do so, I performed ChIP seq experiments using two Pol II CTD specific antibodies: 8WG16, which binds preferentially to unphosphorylated heptapeptides of the Rpb1 CTD, and 4H8, which preferentially binds to the phospho-Ser 5 (YSPTSpPS) form of the CTD. The ChIP profiles obtained with both anti-Pol II antibodies were very similar; >95% peaks overlapped. Notably, most genes occupied by Ell1, Eaf1 and Ebp1 were also occupied by Pol II as shown in both 4H8 (Figure 3.4) and 8WG16 ChIP-seq datasets (data not shown).

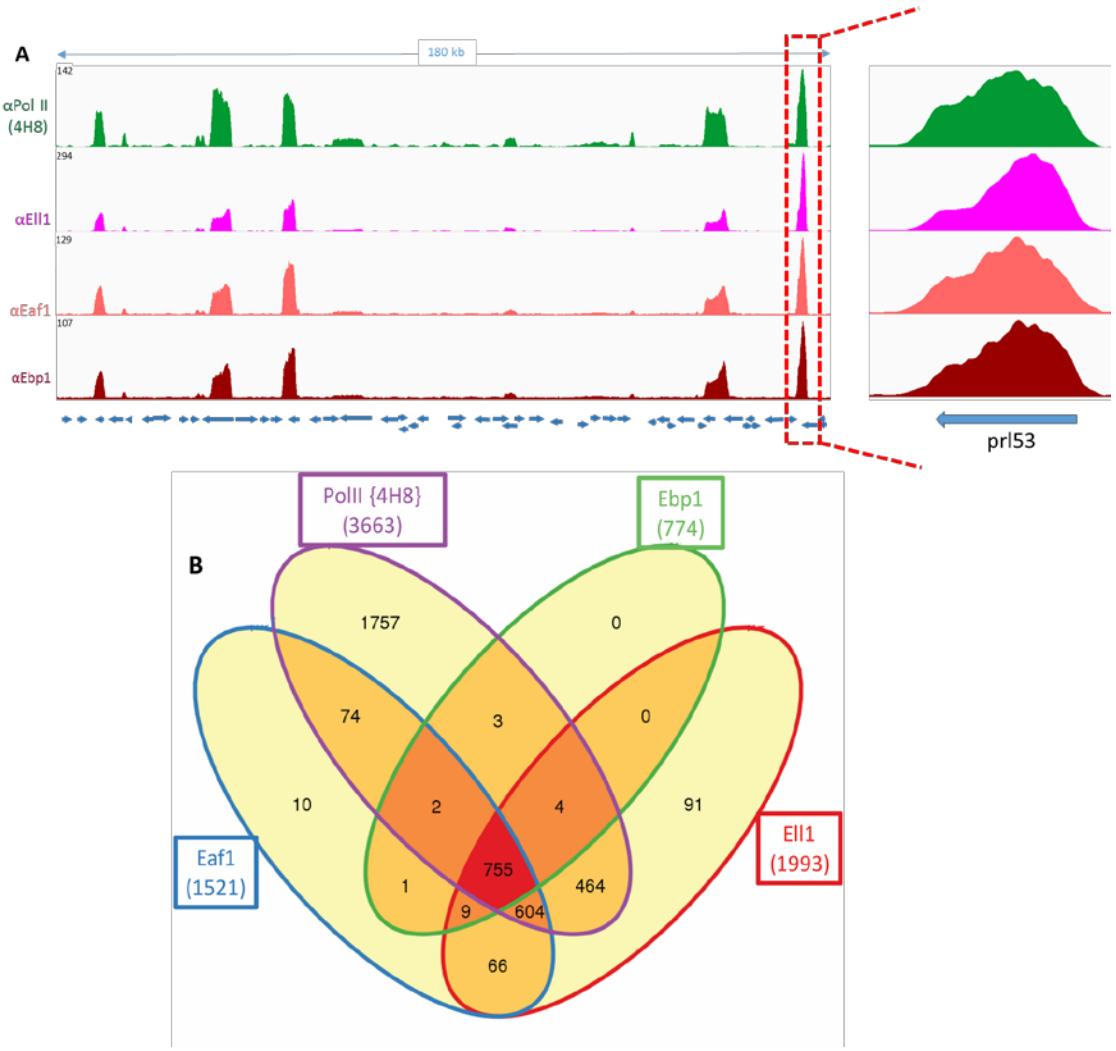


Figure 3.4 Genes enriched by Ell1, Eaf1 and Ebp1 also have high Pol II

A) IGV screenshot of 180kb region in *S. pombe*. ChIP-seq tracks corresponding to Pol II (4H8, shown in green), Ell1 (purple), Eaf1 (orange) and Ebp1 (brown) in Wildtype(WT) strain are shown. The arrows at the bottom represent gene directions. (B) Venn diagram showing overlap between Pol II-, Ell1-, Eaf1-, and Ebp1-occupied genes in wild type *S. pombe*. ChIP peaks corresponding to anti-4H8, Ell1, Eaf1 and Ebp1 were assigned to the closest gene and the four colors represent the number of genes with significant enrichment (50>FC>5, q-value<0.01)

We also checked whether the genes highly enriched in Pol II also have higher levels of “ELL complex”. As shown in Figure 3.5, we observed a high correlation between the amount of Pol II and the amount of Ell1, Eaf1 and Ebp1 present on a gene. There were also a collection of genes that had more ELL complex associated with them than expected based on the Pol II levels. We analyzed these genes and found that Ace2-regulated genes were among the outliers and GO / pathway analysis indicates they are enriched among the outliers. Ace2 and a collection of genes regulated by Ace2, including Adg1, Eng1, and Agn1, show higher ratio of ell1, eaf1, and ebp1 to pol II than was found at many other genes in ChIP-seq. As will be discussed in the next chapter (Chapter 4), we also see that expression of the genes encoding Ace2, Agn1, Eng1 is consistently down-regulated in ell1 and eaf1 mutants, and trend down-ward in ebp1 mutant. Ace2 is a transcription factor that is important for the activation of a cell separation program that results in the dissolution of the septum assembled during cytokinesis between the 2 daughter cells, allowing them to become independent entities. It is involved in the activation of a number of genes like eng1 and atgn1 that encode the hydrolytic enzymes responsible for septum degradation. These observations might suggest that Ell1 and its associated factors Eaf1 and Ebp1 may have a role to play in Ace2-dependent regulatory processes.

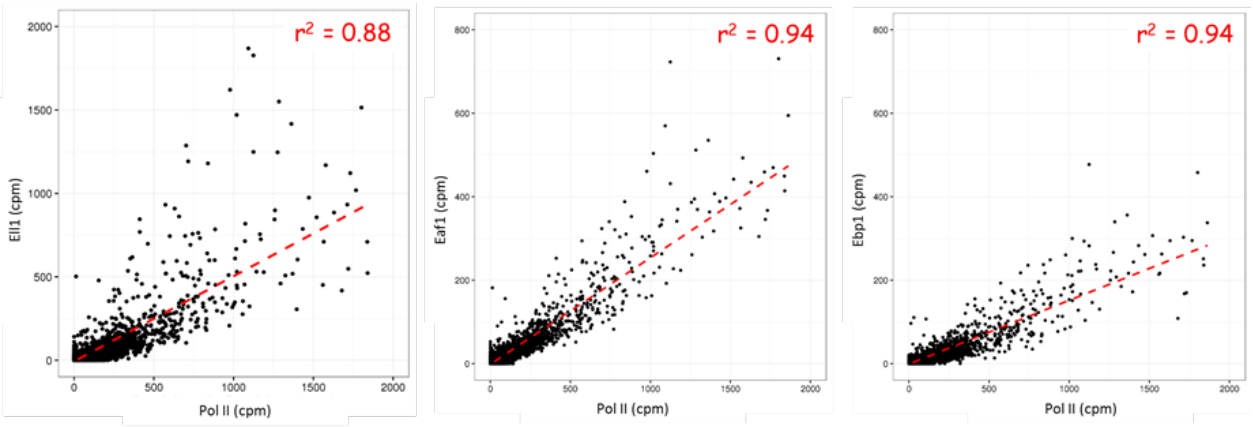


Figure 3.5 Ell1, Eaf1 and Ebp1 occupancy correlate with pol II occupancy

The graph is a plot of read counts per million reads (cpm) obtained in Ell1, Eaf1 or Ebp1 ChIP seq on the Y axis vs the cpm from Pol II ChIP (4H8). r^2 denotes the correlation factor for each graph.

3.6 Ell1, Eaf1 and Ebp1 occupancy correlate with Cdk9 occupancy

In higher eukaryotes, ELL and EAF proteins can be recruited to genes as a part of a larger ELL-containing super elongation complex that also contains AF9 and AF4 family members and P-TEFb. P-TEFb is a highly conserved kinase-cyclin pair composed of CycT and Cdk9, encoded by the *pch1+* and c*dk9+* genes in *S. pombe*. Having observed that Ell1, Eaf1, and Ebp1 are co-recruited to genes in *S. pombe*, we wished to determine whether these Ell complex components are co-recruited with P-TEFb. To do so, we compared our ChIP-seq datasets to a published Cdk9 ChIP-chip dataset (Coudreuse et al., 2010a). As shown in Figure 3.6, we find that Ell1, Eaf1 and Ebp1 occupancy correlate remarkably well with P-TEFb (Cdk9) occupancy. Majority of the genes have the presence of both Cdk9 and Ell1 proteins, as can be seen in the Venn diagram (Figure 3.6B).

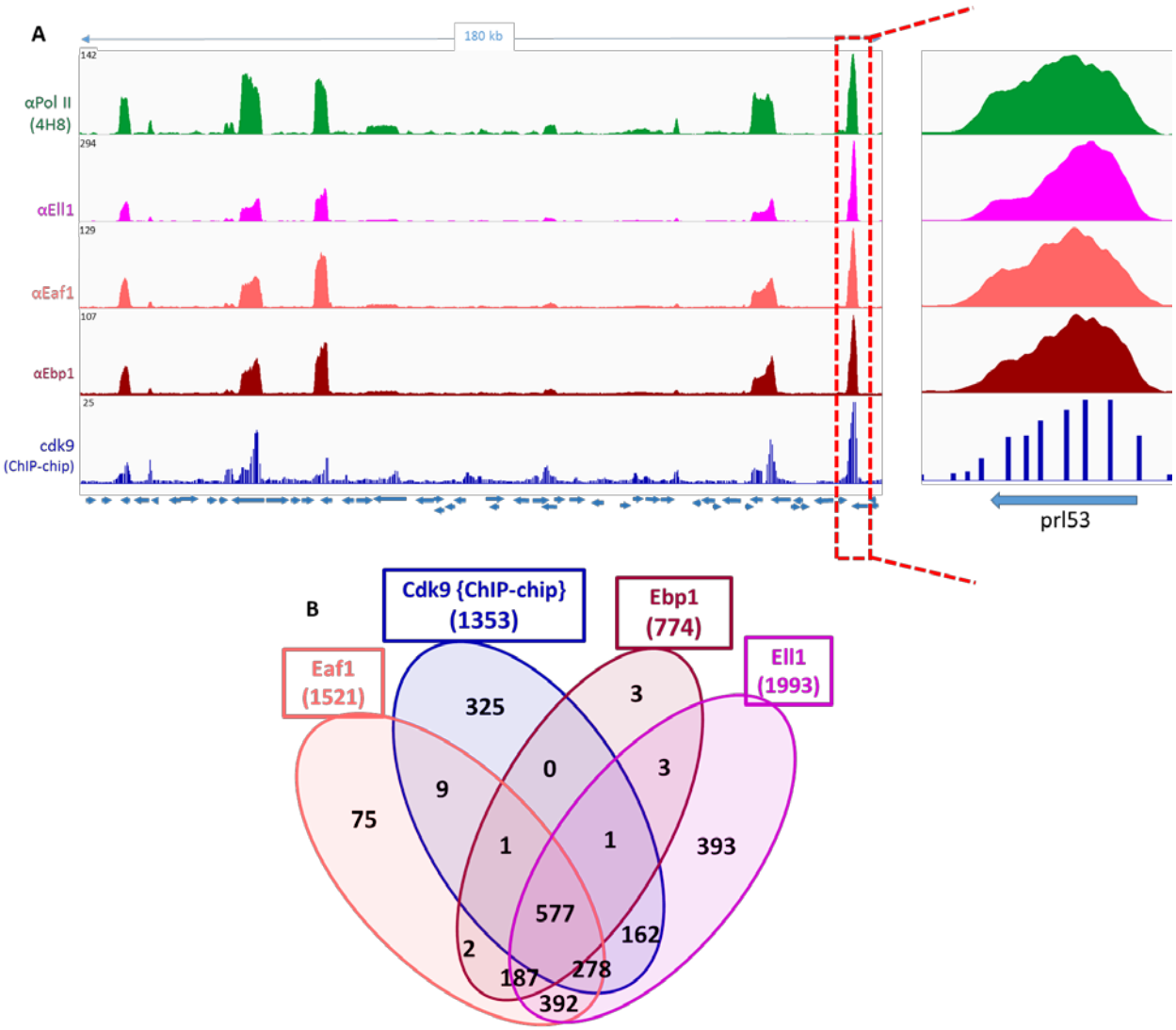


Figure 3.6 Ell1, Eaf1 and Ebp1 occupancy correlate with Cdk9 occupancy

A) IGV screenshot of 180kb region in *S. pombe*. ChIP-seq tracks corresponding to Pol II (4H8, shown in green), Ell1 (purple), Eaf1 (orange), Ebp1 (brown) and Cdk9 (blue) in Wildtype(WT) strain are shown. The arrows at the bottom represent gene directions. (B) Venn diagram showing overlap between Cdk9-, Ell1-, Eaf1-, and Ebp1-occupied genes in wild type *S. pombe*. ChIP peaks corresponding to anti-Cdk9, Ell1, Eaf1 and Ebp1 were assigned to the closest gene and the four colors represent the number of genes with significant enrichment (50>FC>5, q-value<0.01). Cdk9 ChI-microarray data published by (Coudreuse et al., 2010b) was used for analysis.

3.7 Ebp1 does not stimulate elongation by Ell1/Eaf1 *in vitro*

The Ell1/Eaf1 complex has been shown previously to stimulate pol II elongation in vitro (Banks et al., 2007). Since the Ebp1 protein forms a complex with Ell1/Eaf1 and is co-recruited to genes with Ell1/Eaf1, we wished to determine whether it could modulate Ell1/Eaf1 elongation activity. To do so, we tested the effect of Ebp1 protein on the rate of transcription elongation by ternary Pol II elongation complexes assembled on a DNA-RNA scaffold (Kellinger et al., 2012; Walmacq et al., 2009). For these experiments, Flag-Ell1 and Flag-Eaf1 were co-expressed in insect cells and immunopurified on Flag agarose; Flag-tagged Ell1 and Eaf1 were present in the purified fraction in a near equimolar ratio. Flag-Ebp1 was expressed and immunpurified separately (Figure 3.7).

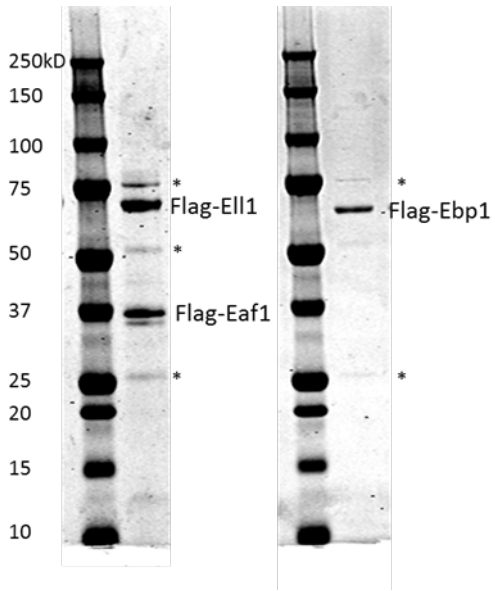


Figure 3.7 Purification of recombinant Flag-tagged Ell1/Eaf1 and Flag-tagged Ebp1

Coomassie blue-stained SDS–PAGE analysis (4–20% gradient gel) of purified recombinant Flag-tagged Ell1/Eaf1 and Flag-tagged Ebp1 from insect cells. Flag-tagged Ell1 and Eaf1 were present in the purified fraction in a near equimolar ratio. The asterisks show nonspecific bands.

As outlined in Figure 3.8, panel A and B, I first assembled an RNA: template strand hybrid and then incubated it with purified *S. pombe* pol II. Biotinylated non-template DNA strand was then added to form the ternary complex, and the mixture was incubated with streptavidin beads. Unbound template strand and Polymerase were then washed off. Transcription was initiated by the addition of ATP and [α-32P] UTP to the reaction mixture and incubated to allow accumulation of pol II ternary elongation complexes containing radioactively labelled, 23 nucleotide long transcripts (Figure 3.8C, lane1). I then chased nascent transcripts into longer products by the addition of ATP, GTP, CTP and UTP, in the presence or absence of Ell1/Eaf1 and Ebp1. In the absence of Ell1/Eaf1, labeled 23-mers were rapidly elongated to 24 nucleotide transcripts, but further elongation proceeded very slowly (Fig. 3.8C, lanes 2-4). As expected, the Ell1/Eaf1 complex stimulated the rate of transcription elongation by its *S. pombe* pol II, as detected by an increase in the rate at which radioactively labelled 24 nucleotide transcripts were chased into longer products when reactions included the Ell1/Eaf1 complex (Figure 3.8C, compare lanes 2-4 to lanes 5-7). We observed little effect on rates of transcript elongation after adding Ebp1 protein to reactions with (lanes 8-16) or without (lanes 17-19) Ell1/Eaf1, suggesting that Ebp1 does not detectably alter Ell1/Eaf1s' ability to stimulate Pol II elongation *in vitro*. The quantification of the assay is shown in the graph in Figure 3.8D.

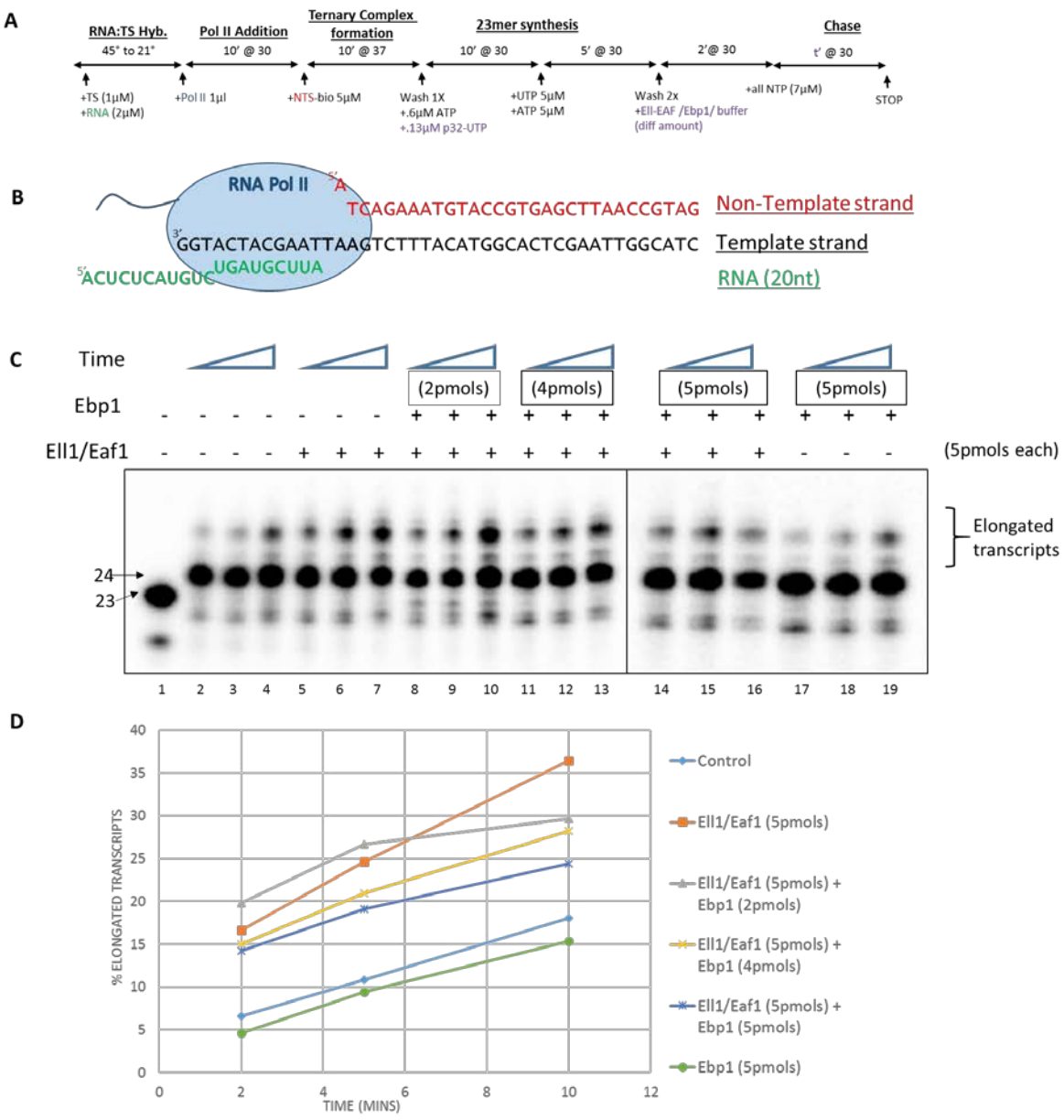


Figure 3.8 Ell1/Eaf1 complex stimulates elongation by *S. pombe* Pol II, but Ebp1 does not

A) Experimental design for ternary complex assembly and transcription. Ternary complexes were formed. Transcripts were labeled by transcription in the presence of 0.6µM ATP and 0.13µM [α-32P]UTP (400 mCi/mmol). After a 10-min incubation, 5µM unlabeled ATP and 5µM UTP were added to ensure transcripts were extended to 23 nt. This 23mer was washed and then incubated in the presence of Ell1/Eaf1 or buffer with/without Ebp1 for 2, 5 or 10 mins of elongation in the presence of 7 µM ATP, CTP, UTP, and GTP. (B) A diagramatic representation of the ternary complex with sequence of the template, RNA and nontemplate strands of the scaffold is shown. Addition of only UTP and ATP would form 23nt RNAs (C) Assays contained either only buffer (lanes 2-4), or 5 pmoles of Ell1/Eaf1 with varying amounts of Ebp1 as shown (lanes 5-16). Reactions shown in lanes 17-19 contained 5 pmoles of Ebp1 without Ell1/Eaf1. Triangles in the figure indicate increasing reaction times of 2, 5 or 10 minutes. (D) The graph shows the quantification of assays.

3.8 Discussion

Ell1 and Eaf1 in *S. pombe* act as elongation factors that share functional similarities with the ELL-EAF complex found in higher eukaryotes. In contrast, *S. cerevisiae* lacks the ELL/EAF homologs. *S. cerevisiae* also lacks the apparatus needed for a number of transcription-linked events, including many splicing processes and RNAi (Kaufer and Potashkin, 2000). That these processes appear to have co-evolved raises the possibility that Ell1 and Eaf1 may have functions *in vivo* related to RNA processing or gene silencing.

We performed MudPIT mass spectrometry analyses and found that both Ell1 and Eaf1 interact not only with each other but also copurify with a previously uncharacterized protein designated as a sequence orphan in Pombase and encoded by the gene *SPAC6G9.15c* (Table 3.1 and 3.2). Analyzing the sequence of this uncharacterized protein by Psi-blast gives regions of possible faint homology to the AF4 family of protein (data not shown), suggesting a possibility that the protein encoded by *SPAC6G9.15c* gene may function as an AF4-Like protein in *S. pombe*. Since this protein binds to *S. pombe* Ell1 we annotated it as an ELL binding protein 1 "Ebp1".

By coexpression of recombinant proteins in insect cells, I have shown that Elll1 can interact directly with both Eaf1 and the gene product of *SPAC6G9.15c* (Ebp1), and most likely forms a bridge between Eaf1 and Ebp1 (Figure 3.1 and 3.2). ChIP-seq experiments using antibodies raised again Ell1, Eaf1 and Ebp1 indicate that these three proteins are colocalized on the *S. pombe* genome, and the pattern of their binding is very similar (Figure 3.3). Since Ell1 binds to Pol II *in vitro* (Banks et al, 2008), we compared Ell1, Eaf1, and Ebp1 occupancy with that of Pol II and observed a high correlation between the

two (Figure 3.4). Gene having the highest Pol II occupancy also had very high 'ELL complex' occupancy (Figure 3.5). Nevertheless, there were many regions with detectable Pol II occupancy where ELL complex components were not enriched. Whether there are many Pol II-occupied genes that lack the ELL complex or whether the apparent difference in the number of genes occupied by Pol II compared to ELL complex is due to differences in antibody affinity remains to be determined.

Since in higher eukaryotes, ELL, EAF and AF4 are part of a larger SEC complex that also contains P-TEFb, we compared our data to a published Cdk9 ChIP-chip dataset and found that Cdk9 occupancy correlates remarkably well with that of the ELL complex and Pol II (Figure 3.6), raising the intriguing possibility that there might be a rudimentary super elongation complex in *S. pombe*.

Finally, I asked whether this new Ell1/Eaf1 interacting protein would have any effect on Ell1/Eaf1s' ability to stimulate Pol II elongation. Under the conditions used in our assays, we detected no Ebp1 stimulation of Pol II elongation rate, either in the presence or absence of Ell1/Eaf1. In fact, if Ebp1 had any effect on Pol II elongation, it was slightly inhibitory. At this point we cannot distinguish between the possibilities that Ebp1 has a weak inhibitory activity or that the purified Ebp1 fraction contained an inhibitory contaminant. Finally, I note that Ebp1 protein's inability to stimulate Pol II elongation may not come as a surprise, since in "bigger eukaryotes," members of the AF4 family of proteins are predicted to act as a scaffolds that link ELL/EAF to P-TEFb (which does not alter elongation in the absence of additional negatively acting transcription factors) (Chou et al., 2013; Lu et al., 2016), and purified ELL-containing complexes with or

without AF4 family proteins stimulate Pol II transcription to similar extents (Biswas et al., 2011).

In the next chapter, I will focus on the effect of mutation or deletion in these genes and how they affect gene regulation.

CHAPTER 4: Consequences of mutation in 'ELL complex'

4.1 Introduction

In the previous chapter, I demonstrated that by mass spectrometry we identified a previously uncharacterized protein, encoded by the *SPAC6G9.15c* gene, as an Ell1/Eaf1 interacting protein, Ebp1. By ChIP-seq, I showed that this Ebp1 protein colocalizes with Ell1, Eaf1, and Cdk9 at genes with high pol II occupancy. Although I observed in biochemical experiments using purified recombinant proteins that the Ebp1 protein does not detectably alter Ell1/Eaf1s' ability to stimulate Pol II elongation *in vitro*, we hypothesize that it contributes to Ell1/Eaf1 function in cells.

In an effort to define *in vivo* functions of Ell1, Eaf1, and Ebp1, I generated *S. pombe* strains deleted for the genes encoding each protein and performed various phenotypic analyses using the resulting *ell1Δ*, *eaf1Δ*, and *ebp1Δ* strains. Consistent with the possibility that Ell1 and/or the Ell1 complex functions in elongation control, I found that deletion of *ell1+*, but not *eaf1+* or *ebp1+*, caused sensitivity to mycophenolic acid (MPA), a phenotype shared with a number of other transcription elongation factors (Reines, 2003). In further experiments, I found that lack of Ell1 complex components had minimal consequences for transcript accumulation or Pol II distribution genome-wide under the growth conditions used. I did, however, obtain evidence that some transcripts in subtelomeric regions may be de-repressed upon loss of *ell1+*, raising the possibility that Ell1 might contribute to heterochromatin formation or maintenance. Consistent with this possibility, I observed a significant decrease in H3K9 methylation marks in subtelomeric regions of *ell1Δ* cells.

4.2 Sensitivity of *ell1Δ* strain to Mycophenolic acid (MPA)

Yeast strains carrying mutations in genes encoding several proteins implicated in regulation of transcription elongation grow slowly in the presence of the nucleotide-depleting drug 6-azauracil and mycophenolic acid. Previous data from our lab shows that *ell1+* deletion in *S. pombe* causes a 6-azauracil sensitivity (Banks et al., 2007). To determine whether *ell1*, *eaf1* or *ebp1* mutants are sensitive to Mycophenolic acid, I grew strains lacking Ell1, Eaf1 and Ebp1 proteins on plates containing the drug. Although the *eaf1Δ* and *ebp1Δ* strains appeared to grow as well as the wild type *S. pombe* (data not shown), the *ell1Δ* strain exhibited a mild sensitivity to 30 µg/ml mycophenolic acid (Figure 4.1).

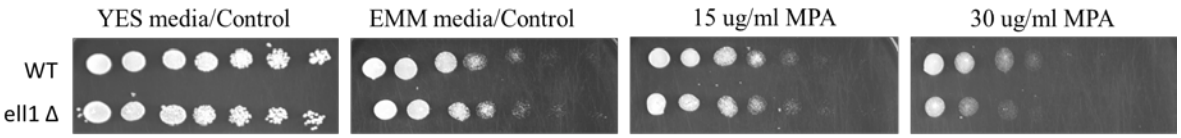


Figure 4.1 Mycophenolic acid sensitivity of *ell1Δ* strain.

The parental Pem2 and *ell1Δ* strains were grown to mid log phase in rich media, washed in 1X PBS, and resuspended in H_2O at a density of 1 x 10^8 cells/ml. 5 µl of 3-fold serial dilutions of cells were spotted onto EMM plates supplemented with adenine, histidine and leucine (225 µg/ml) with or without the indicated concentrations of mycophenolic acid.

4.3 Effect of *ell1+*, *eaf1+* or *ebp1+* deletion on DNA damage.

Transcription-coupled DNA repair pathways enable lesions that block transcription to be repaired more quickly than similar lesions in other parts of the genome. In mammalian cells, ELL has been implicated in restart of transcription after transcription-coupled repair (Mourgues et al., 2013b), and cells depleted of ELL were seen to be hyper-sensitive to UV-irradiation. We therefore wished to determine whether deletion of *ell1+*, *eaf1+* and *ebp1+* would render *S. pombe* hyper-sensitive to DNA damaging agents. To do so, I performed spot assays to compare growth of wild type *S. pombe* to growth of *ell1Δ*, *eaf1Δ* and *ebp1Δ* strains in the presence of various DNA damaging agents, including UV irradiation, Methyl methanesulfonate (MMS) and Hydrogen peroxide(H_2O_2). As shown in Figure 4.2, we observed no reproducible difference between growth of wild type and mutant strains over 3 days of growth after spotting or plating. Recent studies (Dabas et al., 2018; Sweta et al., 2017) have reported a role for Ell1 and Eaf1 under DNA damage stress, conferring upon the mutant cells a sensitivity when grown with MMS or hydroxyurea. We, however, see no significant changes in growth upon deletion of *ell1+*, *eaf1+* and *ebp1+* when grown under optimal conditions on plates containing DNA damage agents.

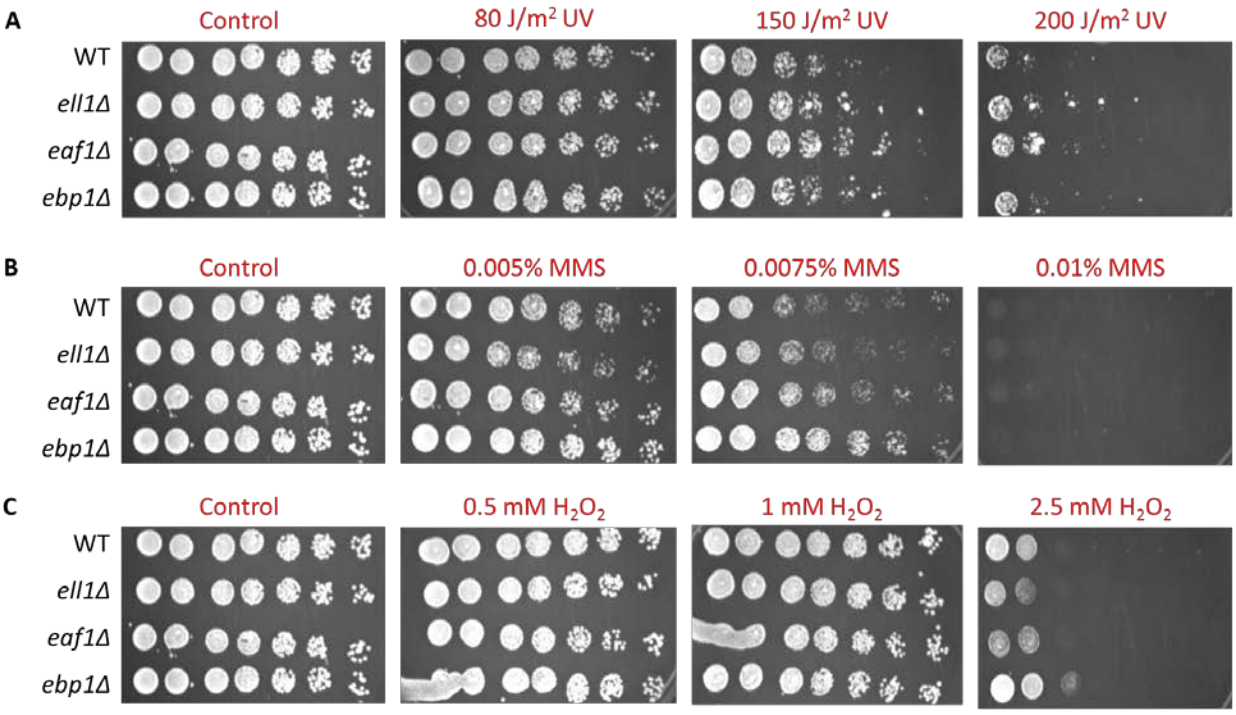


Figure 4.2 Effect of deletion of *ell1+*, *eaf1+* or *ebp1+* on sensitivity to DNA damaging agents.

The parental strain Pem2 and the *ell1Δ, eaf1Δ* and *ebp1Δ* strains, were grown to mid log phase in rich media, washed in 1X PBS, and resuspended in H_2O at a density of 1 x 10^8 cells/ml. 5 µl of 3-fold serial dilutions of cells were spotted onto YES plates without or with varying concentration of (B) methyl methanesulfonate, and (C) hydrogen peroxide. (A) For UV damage, cells were first plated on YES plates and then exposed to the indicated dosage of UV. All plates were incubated at 32°C for 3 days and then imaged.

4.4 Effect of deletion of *ell1+*, *eaf1+* or *ebp1+* on stability of ELL complex

We wished to determine whether recruitment of ELL complex components to genes is inter-dependent. To do so, we performed ChIP using antibodies raised against Ell1, Eaf1 and Ebp1 in wild type and mutant strains lacking *ell1+*, *eaf1+* and *ebp1+* genes, and then quantified recruitment to the promoter region and 3'end of the tdh1 (glyceraldehyde-3-phosphate dehydrogenase) gene and, as a control for the specificity of ChIPs, to a distant non-transcribed region (the K region). Based on results of the ChIP-seq experiments, we expect to see a high enrichment of Pol II and ELL complex at the *tdh1+* gene body and promoter, but not at the transcriptionally silent K region. We observed that deletion of *ell1+* gene decreased the occupancy of not just the Ell1 protein at the *tdh1+* gene, but also of the Eaf1 and Ebp1 proteins (Figure 4.3). Similarly, deletion of either *eaf1+* or *ebp1+* also reduced the occupancies of all three ELL complex proteins.

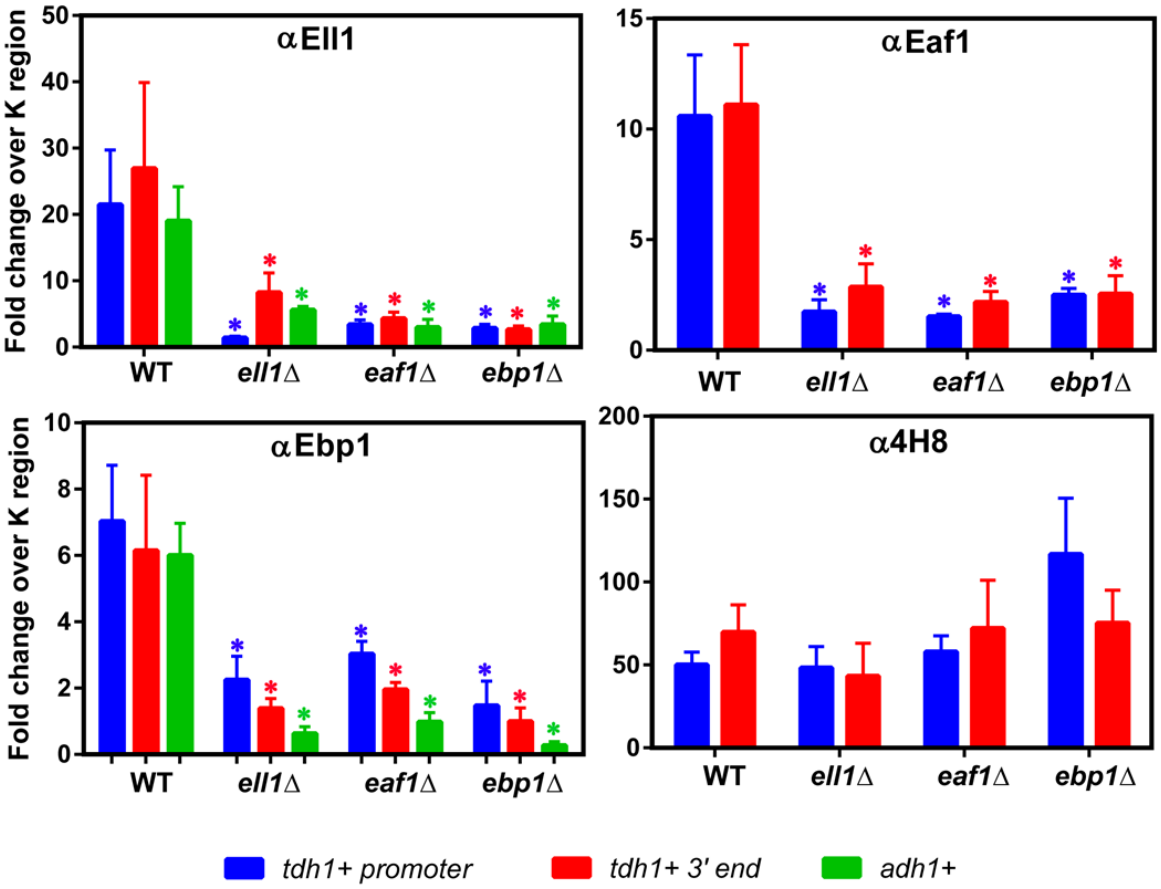


Figure 4.3 Deletion of either one of the proteins decreases Ell1, Eaf1 AND Ebp1 occupancies.

Chromatin immunoprecipitation was performed using antibodies raised against Ell1 (α-Ell1), Eaf1 (α-Eaf1), Ebp1 (α-Ebp1) or Pol II (α-4H8) on wildtype strain and strains having *ell1+*, *eaf1+* and *ebp1+* deletion. Immunoprecipitated DNA was quantified by q-PCR using primers specific to the *tdh1+* gene promoter region (blue), *tdh1+ gene* 3' end (red) and *adh1+* (green). Fold enrichment was calculated over the K region (control region) over the input DNA. The experiment was done in triplicates and the bars represent the standard errors. The "*" represents significant change from WT with $p<0.05$.

These findings suggest either (i) that recruitment of Ell1, Eaf1, and Ebp1 is interdependent or (ii) that deletion of one or more protein(s) of the complex can modulate the expression of the others. To test the latter possibility, we asked whether expression of Ell1 and Eaf1 was reduced in *eaf1Δ* and *ell1Δ* strains respectively. Because

our antibodies could not reliably detect Ell1 or Eaf1 in crude cell lysates against the background of total protein, we performed immunoprecipitation experiments with antibodies against Ell complex components and analysed the immunoprecipitated proteins by western blots. Figure 4.4 shows the results of immunoprecipitations performed using anti-Ell1 or anti-Eaf1 antibodies on equal amount of whole cell lysates from wildtype and *ell1 Δ* strains. In the wild type strains Eaf1 was readily detected in both anti-Ell1 (as it interacts with Ell1) and anti-Eaf1 immunoprecipitates. As expected neither Ell1 nor Eaf1 were detected in anti-Ell1 IPs performed using lysates from the *ell1Δ* strain. Surprisingly, however, much less Eaf1 protein was detected in anti-Eaf1 IPs performed with lysates from the *ell1Δ* strain than from wild type cells, suggesting that the stability or expression of the Eaf1 protein is dependent upon presence of Ell1. Similarly, immunoprecipitation experiments using anti-Ell1 or anti-Eaf1 antibodies were performed on equal whole cell lysates from wildtype and *eaf1Δ* strains. While in the wildtype strain Ell1 can be pulled down through Eaf1, deletion of *eaf1+* made Ell1 undetectable in both anti-Ell1 and anti-Eaf1 IPs, suggesting that expression of Ell1 also depends upon the presence of Eaf1 in the cell.

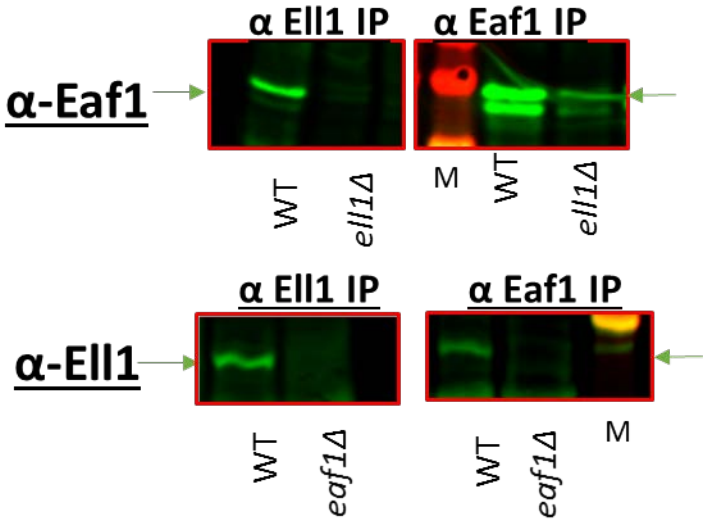


Figure 4.4 Expression of Ell1 and Eaf1 proteins seems to be interdependent.

Immunoprecipitation using α-Ell1 and α-Eaf1 was performed on wildtype strain and strains having *ell1+*, or *eaf1+* deletion. Immunoprecipitated proteins were analyzed by western blotting using anti-Eaf1 and anti-Ell1 antibodies, respectively. The lanes marked 'M' are the molecular size marker lanes. The strong red band in the anti-Eaf1 blot is the 37 kDa marker; the yellow band seen in anti-Ell1blot is 75 kDa.

4.5 Effect of deletion of *ell1+*, *eaf1+* or *ebp1+* on Pol II occupancy

Results from ChIP-seq experiments performed with wild type *S. pombe* suggest that ELL complex is recruited to genes with high Pol II. To determine whether deletion *ell1+*, *eaf1+* and *ebp1+* affects the amount or 5' to 3' distribution of Pol II genome-wide, we performed Pol II ChIP -seq experiments in the mutant strains using 4H8 and 8WG16 antibodies. As shown in the metagene plots in Figure 4.5, deletion of *eaf1+* or *ebp1+* does not detectably alter either the amount or 5'-3' distribution of Pol II ChIP signal, whether we included all genes in the analysis or focused specifically on the most highly expressed genes. On the other hand, the metagene analysis of Pol II distribution in the

ell1Δ strain hints at a possible shift in Pol II distribution towards the 5’ end of the gene. The possible 5' shift in Pol II localization is discernable in metagene plots including data from all expressed genes; however, it becomes more apparent in plots that include only the most highly expressed genes (top 5%, 351 genes). If these findings reflect a real shift, the data is consistent with the possibility that Ell1 is an elongation factor *in vivo* and that its deletion causes Pol II to move slower and hence accumulate toward the 5’ end of the gene.

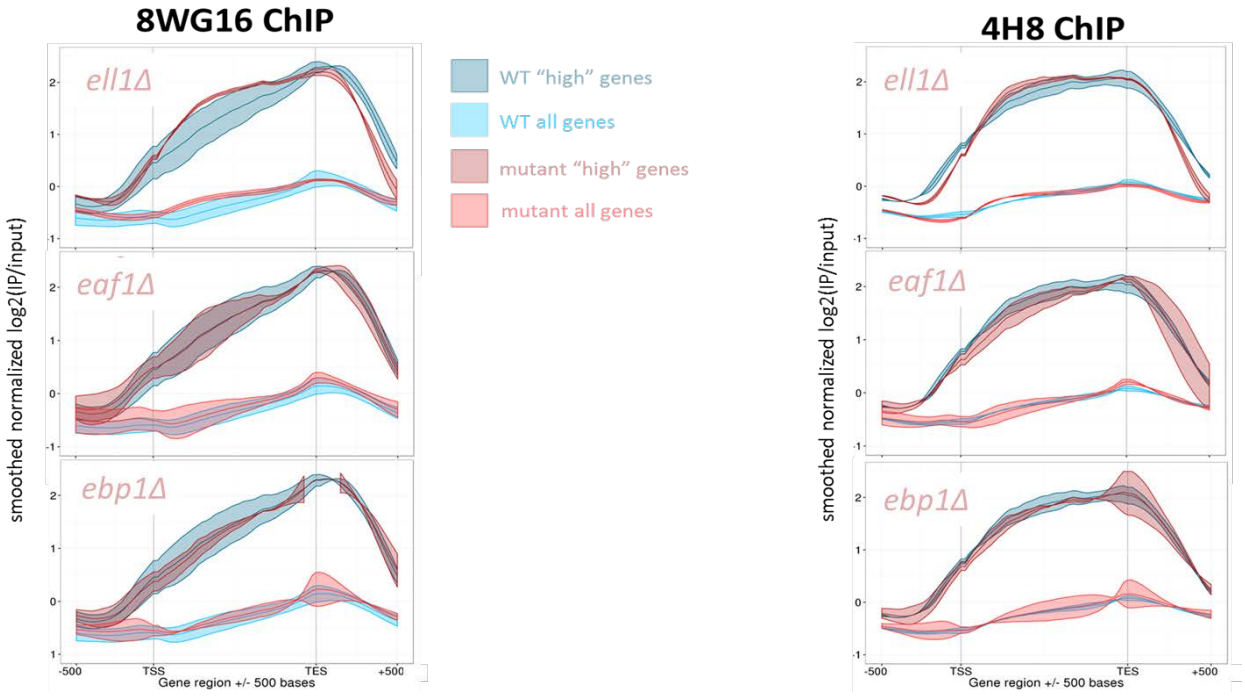


Figure 4.5 Effect of *ell1+*, *eaf1+* or *ebp1+* deletion on Pol II distribution

Chromatin immunoprecipitation using 8WG16 and 4H8 antibodies was performed on wildtype strain and strains having *ell1+*, *eaf1+* or *ebp1+* deletion. Metagene plots of around all the genes having Pol II enrichment made. The blue lines show WT pol II enrichment whereas the red line shows the pol II enrichment in mutant strains as marked. The average normalized reads for “all” genes or the top 5% expressing “high” genes (351 genes) are plotted with the shaded region showing 95 percentile distribution of the average reads in replicates.

4.6 Nascent transcriptome analyses by PRO-seq

Our ChIP-seq analyses provided information about the location of Pol II on genes and raised the possibility that there might be a modest shift in Pol II the 5' to 3' distribution of Pol II across genes. ChIP-seq, however, is a relatively low-resolution method for mapping Pol II occupancy. In addition, it provides no information about the directionality (sense vs antisense) of transcribing Pol II. A recently developed method, Precision Run-On sequencing or PRO-seq (Kwak et al., 2013; Mahat et al., 2016), provides strand-specific information about the precise locations of the 3' end of nascent transcripts. Notably, recent data obtained using PRO-seq identified a previously unrecognized pause in early elongation in *S. pombe* but not in *S. cerevisiae* (Booth et al., 2016); this pause resembles the promoter-proximal pausing in metazoans. To explore the effect of *ell1+* deletion on transcription and to get a base-pair resolution map of transcription elongation in *S. pombe* lacking *ell1+*, we investigated the effects of the mutation on the positions of RNA polymerase active sites genome-wide using Precision Run-On sequencing. If loss of *ell1+* leads to a major defect in release from promoter-proximal pausing, we expect to observe a significant 5' shift in the distribution of reads within the first few hundred bases of the transcription start site. We performed PRO-seq experiments on wildtype and *ell1Δ* strains, in duplicate. The resulting data, summarized in the metagene plots shown in Figure 4.6A, suggests that *ell1+* deletion does not cause major changes in the distribution of polymerase within the first 500 bases of the transcription start site of genes that were separated from the boundaries of neighboring genes on the same strand by at least 1 kb. In further analyses, we specifically selected for genes that may be paused, by calculating the pausing index (mappable reads within

the promoter-proximal region of the gene/mappable reads on gene body), and plotted the average PRO-seq signals around the TSS of these genes in the wildtype and *ell1Δ* strains (Figure 4.6B). As expected, we detect a larger accumulation of polymerase at the 5' ends of these genes as compared to the complete gene set in Figure 4.6A, but again we observe no significant change in polymerase distribution in the wildtype and *ell1Δ* strains.

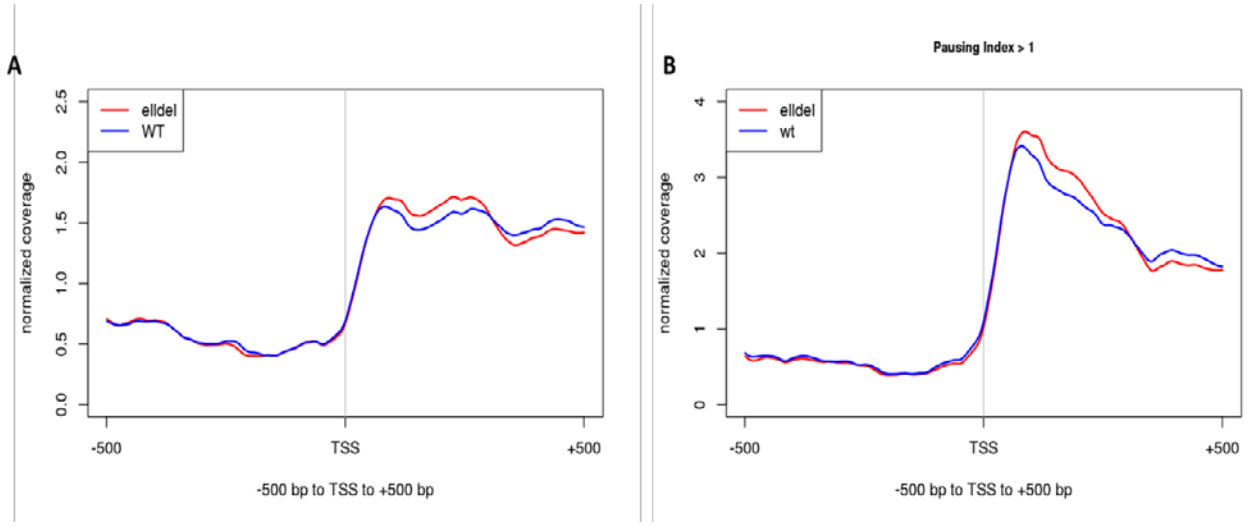


Figure 4.6 5' ends of genes exhibit no major change in Pol II density upon ell1+ deletion.

(A) Average PRO-seq signal in WT (blue) and ell1 Δ (red) strains, around the transcriptional start site (TSS) of active, filtered to include only genes that are longer than 1 kb and separated from the boundaries of neighboring genes on the same strand by at least 1 kb. (B) Average PRO-seq signal around the TSS of paused genes (having pausing index >1, total 960 genes) in WT (blue) and *ell1Δ* (red) strains.

We and others (Chen et al., 2012; Ni et al., 2010) have observed prevalent anti-sense transcription at many genes in *S. pombe*. Because Pro-Seq provides information about the strand specificity of nascent transcripts, we interrogated our dataset to determine whether deletion of *ell1+* would have any effect on the ratio of sense vs antisense transcription. To do so, we calculated the Log2 values of the ratios of the normalized (RPKM) reads on the sense strand to the normalized reads on the antisense

strand and plotted the values for all genes in *ell1Δ* and wildtype strains (Figure 4.7). We find that points corresponding to most of the genes fall on the 'X=Y' line with a slope of 1, suggesting that ratio of sense to antisense transcription of most annotated genes does not change upon *ell1+* deletion.

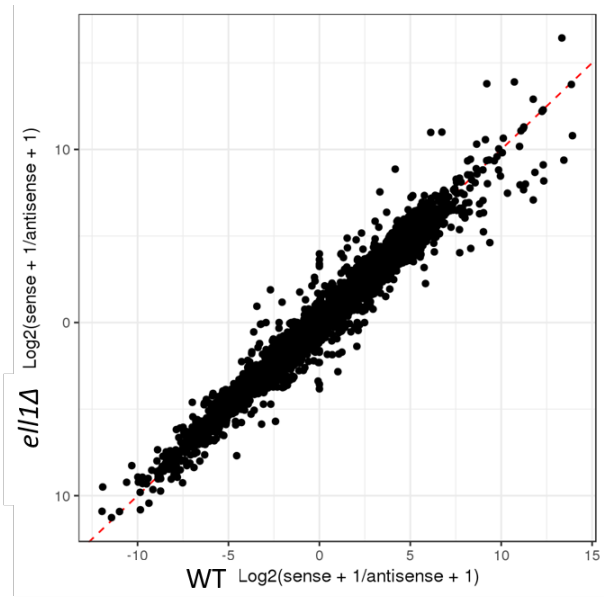


Figure 4.7 Antisense transcription on genes is not affected by ell1+ deletion.

Log2 value of the RPKM ratios on the sense to the antisense strand of individual genes were calculated and plotted for wildtype and *ell1+* deletion strains. (1 was added to both the numerator and denominator to account for strands with reads). The red dotted line denotes the 'X=Y' line with a slope of 1.

4.7 Effect of deleting *ell1+*, *eaf1+* or *ebp1+* on mRNA abundance

To explore the consequences of deleting components of the ELL complex for global gene expression in *S. pombe*, we investigated the effects of deleting *ell1+*, *eaf1+* or *ebp1+* on the abundance of poly-adenylated transcripts. To assay the relative abundance of all expressed transcripts, we prepared strand-specific, poly-A selected RNA-seq

libraries from the wild-type and *ell1Δ*, *eaf1Δ*, and *ebp1Δ* strains of fission yeast. The deletions of each gene in the mutant strains were confirmed by the absence of any reads from the deleted locus. Transcript levels were found to be highly reproducible between biological replicates.

We identified genes that were differentially expressed in the mutant strain compared to wild type using EdgeR. As summarized in Fig. 4.8, we observed that in all three deletion strains tested, only a fairly small number of genes were up or down regulated upon *ell1+*, *eaf1+* or *ebp1+* deletion (data for poly-A selected RNA seq libraries shown in Figure 4.8A). A significant number of these genes overlap in both the upregulated (Figure 4.8B) and the downregulated categories (Figure 4.8C).

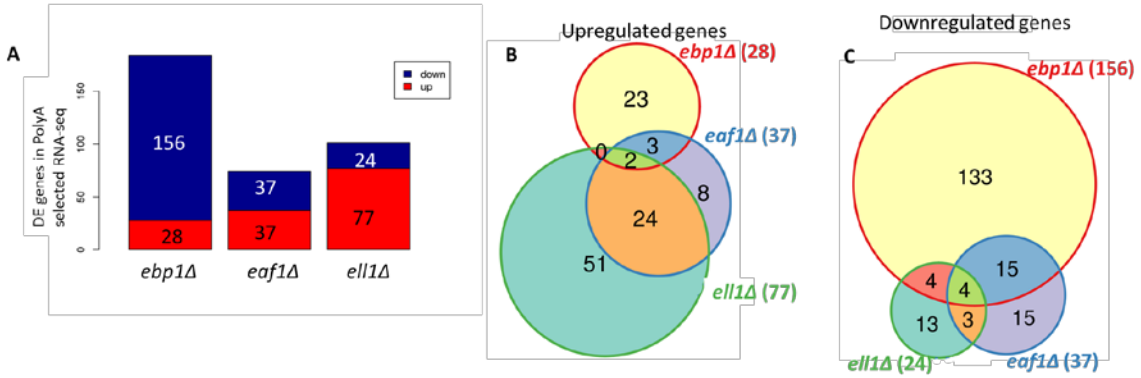


Figure 4.8 Differentially expressed genes in poly A selected RNA seq libraries upon *ell1+*, *eaf1+* or *ebp1+* deletion.

(A) Differentially expressed (DE) genes in the polyA selected RNA-seq analysis of *ebp1Δ*, *eaf1Δ* or *ell1Δ* with respect to the wildtype strain (with 1.5-fold difference or more, P<0.05). The overlap between the genes that were (B) upregulated or (C) downregulated in the three strains tested is shown.

We find that many genes that were up or down regulated upon deletion of one gene do not show up in other mutant strains as differentially expressed. This is probably

because we are using stringent cutoffs to call a gene differentially expressed. Although they do not pass our cutoffs for significant differential expression, most of these genes do exhibit a similar trend in expression change over the wild type strain, with the foldchange and/or the adjusted p-value barely making the cutoffs. Of note, Ace2 and Ace2-regulated genes were also seen to be downregulated upon Ell1 deletion. The *eaf1+* and *ebp1+* deletion also decreased the stable transcripts of these genes modestly.

We also generated 'ribo-depleted RNA' libraries to allow comparison of the expression of non-coding RNAs that may lack poly-A signal and find that the overall trends are similar to what was found in polyA+ group. We again found a very modest effect of depletion of Ell1, Eaf1 or Ebp1 on steady state RNA levels; however, on closer examination, we noticed that some of the genes upregulated upon *ell1+* deletion were located next to one another, and quite close to the right end of chromosome 1, suggesting that their increased expression could be a consequence of their similar chromosomal locations rather than any functional relationship. It should be noted that the total reads of many of the genes in the telomeric and subtelomeric regions are quite low as these genes, located in the heterochromatin region, are expressed at extremely low levels. A group of genes that seems to be upregulated in the subtelomeric region in the *ell1Δ* strain are shown in Figure 4.9. This observation raised the possibility that Ell1 might play a role in the telomeric/subtelomeric region of chromosomes.

GENE	LOG(FC)
SPAC977.03	2.30
SPBPB2B2.19c	2.30
SPAC186.08c	1.77
SPAC186.07c	1.74
SPAC750.05c	1.61
SPAC977.01	1.29
SPAC750.01	1.10
SPAC186.04c	0.84
SPAC186.09	0.57
SPNCRNA.448	0.54
SPNCRNA.61	0.53
SPBPB2B2.18	0.50

Figure 4.9 Subtelomeric genes upregulated in *ell1Δ*

Genes with at least a 1.4-fold (log2 > 0.48) increase in transcript abundance in the *ell1Δ* strain relative to wildtype.

4.7 Altered subtelomeric H3K9 methylation in *ell1Δ* strain

The regulation of heterochromatin is essential for proper chromosome segregation, genomic stability, and cell fate determination. In *S. pombe*, heterochromatin is found at the telomeric, subtelomeric, and peri-centromeric regions and at the mating type loci and is characterized by high levels of H3K9 methylation. To examine the heterochromatin state at the centromeric and the telomeric regions, we performed ChIP-seq experiments in the wild type and *ell1+* deletion strains using an antibody that recognizes the H3K9 dimethylation mark. We also performed ChIP seq with an antibody that recognizes histone H3 to determine whether any observed changes could be due to changes in total H3 occupancy. We detected no effect of *ell1+* deletion on the methylation pattern at the centromeric regions of the three *S. pombe* chromosomes (Figure 4.10).

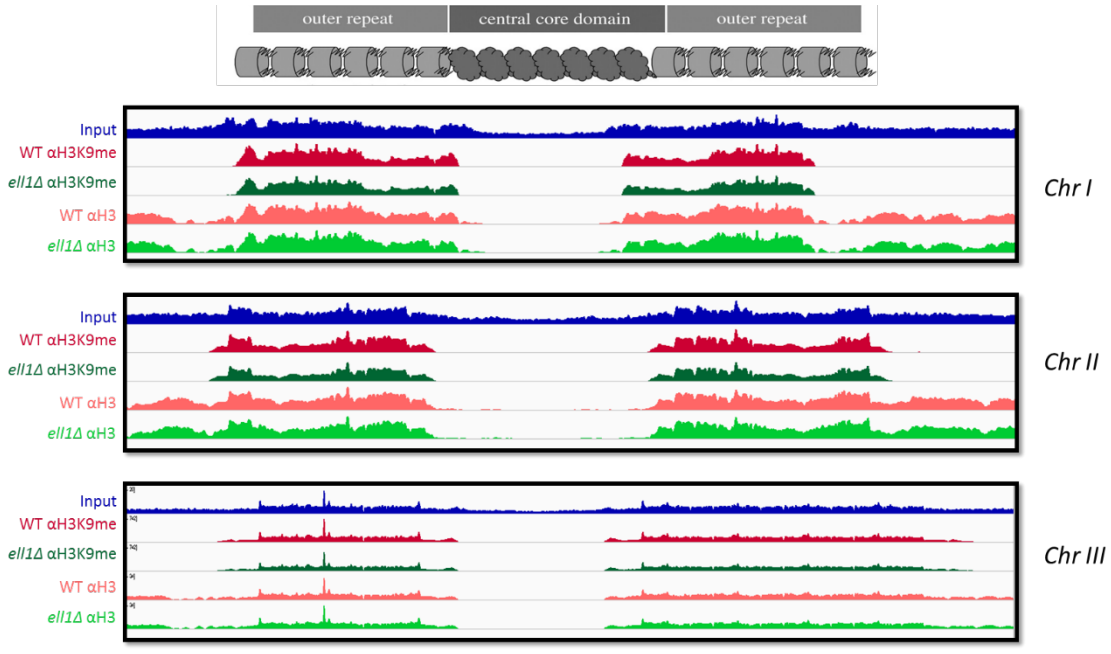


Figure 4.10 Heterochromatin distribution around the centromere is similar in wild type and *ell1Δ* strains.

IGV tracks showing outer centromeric repeat (OTR) and the central core domain of centromeric regions in *S. pombe*. The tracks show H3K9(me2) ChIP and H3 ChIP in the wildtype (red) and *ell1Δ* (green) strains on the three chromosomes of *S. pombe*. An example of an input DNA track is shown in blue.

To examine the effect of *ell1+* deletion on H3K9 methylation in subtelomeric regions, we focused on chromosomes I and II, since interpretation of ChIP-seq data at the ends of chromosome III is complicated by the presence of rDNA repeats. As shown in Figure 4.11, we see altered subtelomeric H3K9 methylation patterns in the *ell1+* deletion strain at the left and right ends of chromosome I and II (Figure 4.11). We calculated H3K9methlyation enrichment over the input DNA and see no change in H3K9 methylation levels immediately proximal to the telomeres (shown in green boxes). In contrast, when we look at the subtelomeric region shown in red boxes, we see a significant decrease in the methylation marks in the *ell1Δ* cells, suggesting that Ell1 may play a role in maintaining the heterochromatin at the subtelomeric regions.

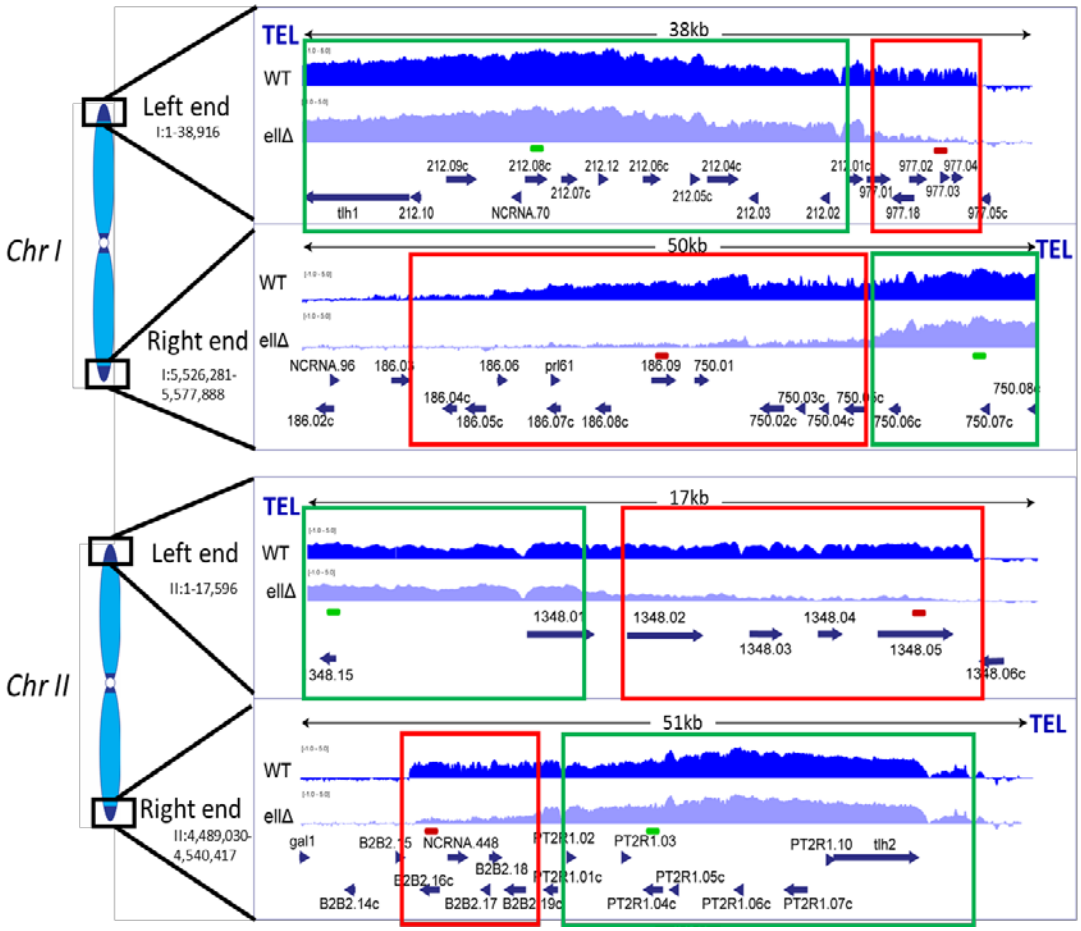


Figure 4.11 Altered subtelomeric H3K9 methylation in *ell1Δ* strain.

IGV browser shots of telomeric regions at the left and right ends of Chromosome I and II in *S. pombe*. The tracks show normalized H3K9(me2) enrichment (ChIP/input DNA) in wildtype and *ell1Δ* strains. The green boxes show region proximal to the telomere. The red boxes show the subtelomeric region with decreased H3K9 methylation in the *ell1+* deletion strain. The genes in the region are shown below the tracks as arrows denoting the directionality of the genes. The green and red bars at each chromosome end show the position of primers used for q-PCR in Section 4.8 and Figure 4.12.

4.8 *eaf1+* or *ebp1+* deletion does not alter subtelomeric H3K9 methylation

To confirm the changes in H3K9 dimethylation mark in the *ell1Δ* strain detected by ChIP-seq and to determine whether deletion of *eaf1+* or *ebp1+* would give rise to the same phenotype, we performed ChIP-qPCR experiments using primers specific for the ends of chromosome I and II for all the mutant strains. The red and green bars denoting the primer positions are shown in Figure 4.11. The 'red' primers are in the subtelomeric region where H3K9 methylation is altered in the *ell1Δ* strain, whereas the 'green' primers are in the regions closer to telomere where the heterochromatin is unaltered upon *ell1+* deletion. We observe that deletion of *ell1+* causes the enrichment of DNA amplified by 'red' primers to go down while the region recognized by 'green' primers remains unaltered. This result is consistent with the ChIP-seq data from section 4.7, as in the *ell1Δ* strain the subtelomeric H3K9 methylation mark goes down, but there is little or no change in the region closer to telomeres. In contrast, deletion of *eaf1+* or *ebp1+* does not decrease the enrichment of H3K9 dimethylation in the subtelomeric region or the region closer to the telomere, suggesting that Ell1 may have roles independent of Eaf1 and Ebp1. We do see an increase in one replicate of eaf1Δ strain, but the control values were also higher, hence we are not focusing on it.

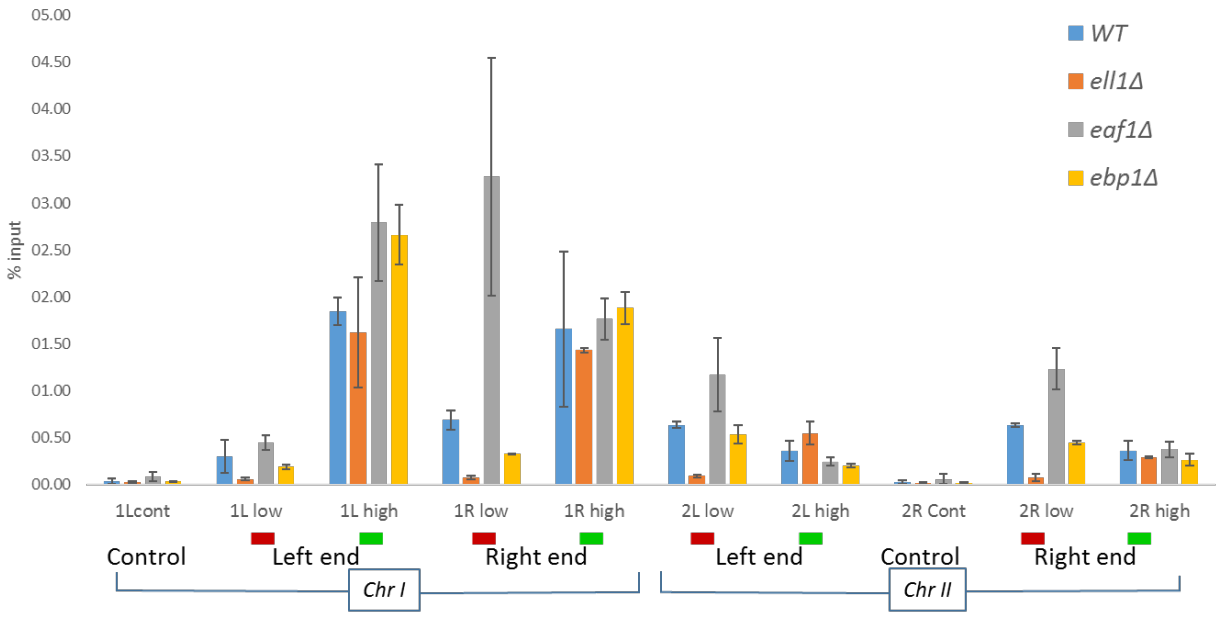


Figure 4.12 Unlike *ell1+* deletion, deletion of *eaf1+* or *ebp1+* does not alter subtelomeric H3K9 methylation.

ChIP qPCR experiments were performed on Wildtype (WT), *ell1Δ, eaf1Δ* and *ebp1Δ* strains using antibodies recognizing the H3K9(me2) mark. The enrichment is calculated as % input DNA. The 'red' and 'green' primers from the left and right ends on Chromosome I and II are used (position as indicated in Figure 4.11) and the control set of primers are in regions in each Chromosome having no H3K9 methylation (from ChIP seq data).

4.9 Discussion

6-Azauracil (6AU) is an inhibitor of IMP dehydrogenase (IMPDH), the rate-limiting enzyme in de novo GTP synthesis and therefore treatment of cells with 6AU results in depletion of intracellular nucleotide pools. Mycophenolic acid is also a specific inhibitor of IMPDH. Mutations in a number of genes encoding proteins implicated in elongation control give rise to 6AU and mycophenolic acid sensitivity. Loss of Ell1 also gives rise to decreased growth rate in the presence of drugs that inhibit nucleotide synthesis, including 6AU (Banks et al., 2007) and MPA, and therefore is consistent with the idea it has a role in transcription elongation.

We also checked the deletion mutants for growth defects when grown with various DNA damage causing agents and found that the mutants grow very similarly to the wildtype strain. Thus, at least under the growth conditions used, the ELL complex does not appear to play a major role in the response to DNA damage. As noted earlier, work from another lab (Dabas et al., 2018; Sweta et al., 2017) reported increased sensitivity to DNA damaging agents in ell1 and eaf1 deletion mutants, including some of the same strains I have used. At present, we do not understand the reason for this discrepancy.

Deletion of *ell1+* did lead to a subtle shift in pol II distribution toward the 5' end of genes, suggesting that polymerase could lag behind upon loss of *ell1+.* This again points towards role of Ell1 in transcription elongation *in vivo*. However, deletion of *eaf1+* or *ebp1+* did not lead to a discernable change in pol II distribution across genes.

Steady state levels of RNA transcripts did not seem to be not greatly affected by deletion of ELL complex components. However, we noted a cluster of genes in the region close to the right arm of Chromosome 1 that seemed to be modestly upregulated upon *ell1+* deletion. Chromosomal location of these genes rather than the functions may be key regulation.

Fission yeast subtelomeric regions can be divided into 2 regions having high H3K9 methylation in the telomere-proximal region and lower H3K9 methylation in the region more distal from the telomeres. It is possible that H3K9methylated subtelomeric heterochromatin protects *S. pombe* subtelomeres from recombination between nearly identical paralogs, while the lower level of H3K9methylation in the more distal region still permits transcription of genes in this region as needed.

It is interesting to note that phosphorylation-site mutations in the CTD of the largest Pol II subunit have also been reported to alter subtelomeric gene expression in pombe (Inada et al., 2016). In addition, these authors observed that manipulation of the $Y^1S^2P^3T^4S^5P^6S^7$ heptapeptide repeat of the CTD by substituting non-phosphorylatable alanines for Ser2 and/or Ser7 led to the alteration of the repressive histone H3 lysine 9 methylation (H3K9me) landscape. Similar phenotypes were observed when cohesin loader mutants were studied in *S. pombe*, suggesting a role for cohesin in formation of heterochromatin domains at the the subtelomeres (Dheur et al., 2011). CTD phosphorylation is mediated by the protein kinase and transcription elongation factor P-TEFb, which is known to interact with ELL as a part of SEC in higher organisms. Studies also report that mutations in cohesin impairs both Pol II transcription initiation at promoters and elongation through the gene body (Mannini et al., 2015). It is interesting to note that the changes in subtelomeric H3 K9 methylation patterns in CTD and cohesin loader mutants were very similar to those we observed in our *ell1Δ* strain. The precise mechanisms by which these mutations affect chromatin structure and gene expression remain to be elucidated. It could be interesting in the future to determine whether mutations in ell1 or cdk9 lead to increased recombination between subtelomeric regions

These 20- to 40-kb regions between the chromosome arm euchromatin and the telomere define a functionally distinct so-called subtelomeric region of their respective chromosomes. Chromosome 3 in *S. pombe* has a different structure than the other two chromosomes in that ribosomal RNA gene repeats exist near the ends of chromosome 3, perhaps explaining the distinct change in H3K9me2 patterns seen in *ell1Δ* strain for chromosomes 1 and 2, but not chromosome 3. It is possible that the relatively low level of H3K9methylated subtelomeric heterochromatin protects *S. pombe* subtelomeres from

recombination between nearly identical paralogs of subtelomeric regions, while still permitting the expression of the corresponding genes when needed.

CHAPTER 5: Synthetic Genetic Array (SGA)

5.1 Introduction

In the previous chapters, I demonstrated that we identified a novel Ell1 and Eaf1 interaction partner, which we have named ELL binding protein 1 or Ebp1. I found that Ebp1 colocalizes with Ell1 and Eaf1 at many genomic loci but does not detectably stimulate the rate of transcription elongation by Ell1/Eaf1 in our *in vitro* assays. Together, we refer to Ell1, Eaf1 and Ebp1 as the ELL complex.

We made gene deletions of each of these genes individually to study the role these genes may be playing in the cell. Despite a reported role for mammalian ELL and EAF in the response to UV irradiation (Mourgues et al., 2013a), *S. pombe* strains lacking ELL complex components showed no hyper-sensitivity to DNA damaging agents under the conditions tested. On the other hand, deletion of the *ell1+* gene gave rise to sensitivity to 6-azauracil and mycophenolic acid, a characteristic of many genes involved in transcription elongation (Riles et al., 2004). The deletion of *ell1+*, *eaf1+* or *ebp1+* did not affect the Pol II occupancy genome-wide and affected expression of a relatively small fraction of genes. Results of RNA seq experiments indicated, however, that among those affected by *ell1+* deletion was a group of up-regulated genes near telomeric regions. We performed ChIP seq experiments using antibody that recognizes H3K9 methylation mark, a characteristic mark for heterochromatin in eukaryotes, and found that upon deletion of *ell1+* the heterochromatin pattern at the sub telomeric region was altered. *eaf1+* or *ebp1+* deletions, on the other hand, did not seem to alter the subtelomeric H3K9

methylation mark, suggesting Ell1 may have an independent role in regulating heterochromatin at the subtelomeric regions in fission yeast.

In this section, we exploit yeast genetics in an effort to gain additional insight into the roles of Ell1, Eaf1 and Ebp1 in cells. Our aim is to find genetic interactions of these three genes in *S. pombe* using a high throughput approach. Large-scale genetic interaction mapping can be performed using synthetic genetic array (SGA) analysis, a method that offers an efficient approach for the systematic construction of double mutants and enables a global analysis of synthetic genetic interactions (Roguev et al., 2007a). In a typical SGA screen in *S. pombe*, a query mutation is crossed to an ordered array of ~5000 viable gene deletion mutants such that meiotic progeny harboring both mutations can be scored for fitness defects. Estimating the fitness of the two single mutants and their corresponding double mutant gives a quantitative measurement of genetic interactions, distinguishing negative (synthetic lethal) and positive interactions. Proteins that act separately but have overlapping important functions show negative genetic interaction such that loss of both genes results in severe growth defect. Positive genetic interactions refer to double mutants with a less severe fitness defect than expected and include interactions such as epistasis and suppression

We performed the screen in triplicate for all three of our genes, *ell1*, *eaf1* and *ebp1*. We focused our analysis on the identification of synthetic lethal or sick interactions, in which a combination of mutations in two genes results in cell death or reduced fitness. We found multiple genes that interact with genetically with *ell1*, *eaf1* and *ebp1*. A second group of genes showed significant interactions with only *ell1*. We focused on the most interesting interactions from each group of these genes.

Brl1, an E3 ubiquitin ligase enzyme catalyzes the mono-ubiquitylation of histone H2B at a conserved site in the carboxyl-terminus (H2Bub1) throughout coding regions of genes. *brl1Δ* was found to have a strong genetic interaction with *ell1Δ*, *eaf1Δ* and *ebp1Δ*. H2Bub1 plays a central role in the interplay between chromatin and the RNAPII elongation complex. Formation of H2Bub1 on transcribed chromatin also requires PAF and H2Bub1 promotes co-transcriptional generation of H3K4me by methyltransferase Set1. Reports suggest interdependence of Cdk9-mediated Spt5 phosphorylation and H2Bub1, a co-regulation governed by positive feedback. We explore the interaction between ELL complex and Brl1 further in this section.

It was also interesting to note that we got many genes that are involved in maintaining the heterochromatin at centromeres and telomeres in our *ell1Δ* SGA screen. Our previous data from RNA seq and H3K9me2 ChIP-seq already suggest that Ell1 may have a role formation or maintenance of H3K9 methylation at subtelomeric regions in pombe. Genetic interaction of *ell1Δ* with genes involved in heterochromatin formation further supports this idea.

This section first explains the strategy of performing SGA and then discusses some of the interesting negative genetic interaction we found in our screens.

5.2 Strain construction and synthetic genetic array

SGA analysis requires a relatively simple set up that involves generating the query strain of interest and crossing it to an array of all the viable deletion mutant strains. Through a series of several replica-plating steps, the double mutants are selected and

scored for growth. Each step of the procedure is described in detail below (also shown in Figure 5.1A).

Query strain construction:

The genes of interest were individually deleted using PCR based techniques to generate *ell1Δ*, *eaf1Δ* or *ebp1Δ*, containing a gene deletion marker *natMX4*. *natMX4* provides the cells resistance against the drug nourseothricin, commonly referred to as clonNAT. The strain background used for generation of query strain was Pem2 and followed the protocol as described in (Roguev et al., 2007a). The h^- Pem2 cell contains cyh^S in the h^- locus and a cycloheximide-resistance (cyh^R) allele at the endogenous *rpl42* locus. This configuration mimics a cyh^S- cyh^R diploid and hence is sensitive to cycloheximide. After mating and sporulation, growth in the presence of cycloheximide serves as both anti-diploid and mating-type selection as it allows selection for only the meiotic h^+ haploid and eliminating any remaining diploids or unmated parent cells (shown in Figure 5.1B). The *S. pombe* deletion array was obtained from Bioneer Inc. having all viable gene deletions *in* h^+ *ade6-M210 ura4-D18 leu1-32* genotype with *KanMX6* marker providing resistance to G418.

SGA procedure:

The array of G418 resistant mutants were pinned onto an agar plate such that each plate has 384 clones. Position of each deletion was noted. Similarly, the query strain was spotted onto an agar plate (384 spots/ plate). The plates were incubated at 32°C for 2 days for the cells to grow. After 2 days, 384 clones were robotically pinned onto SPAS plate (minimal media that causes sporulation), followed by the pinning of query strain on top of it and finally 384 spots of water were put on the plate to facilitate sporulation. The

plates were kept at room temperature for ~5 days or until tetrads can be seen under microscope. The spores were printed onto YES plates containing G418 and cycloheximide (100µg/ml) as quadruplicate spots (1530 spots/plate). Cells were grown for 2 days to select for h^- mutants. For further selection, the colonies were replica printed onto YES plates containing G418 and cycloheximide again and grown for 2 days. Finally, the colonies were replica printed onto 'Control' plates (YES+G418+cycloheximide) and 'Test' plates (YES+G418+clonNAT+cycloheximide). Only the double mutants would grow on the Test plates as it contains both G418 and clonNAT, whereas, on the control plates containing only G418 both the single and double mutants survive.

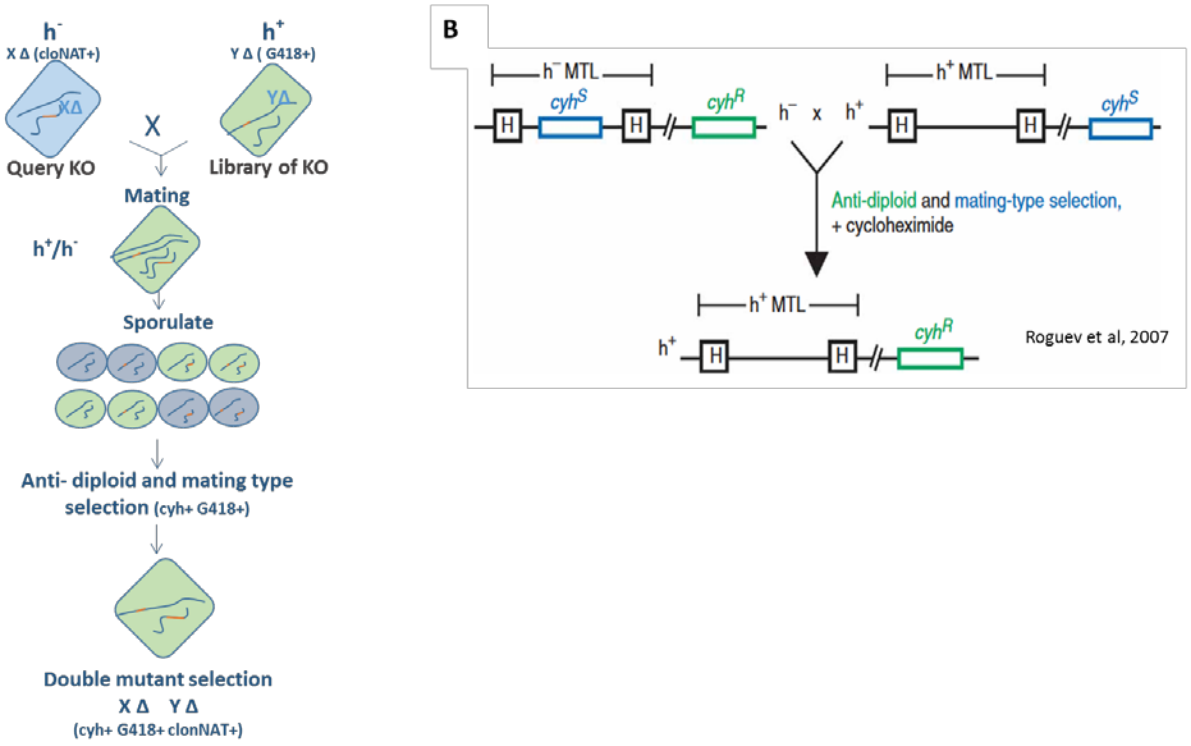


Figure 5.1 SGA methodology using Pem2 strategy.

(A) Two haploid strain carrying differently marked mutation (X and Y) are mated and following meiosis and sporulation, haploid progenies are produced. Anti-diploid selection, mating-type selection and double-mutant selection are then applied. (B) PEM-2 strategy: a cyh^S allele is expressed from within the mating type locus of h^- cells while a cyh^R allele is expressed from the endogenous locus, creating a genotype that mimics a heterozygous diploid. After meiosis, the only cells able to survive on medium containing cycloheximide are haploids from the opposite mating type (h^+).

Image processing and quantitative scoring using colony size-based fitness measurements:

Double mutant array plates are photographed in a light-controlled chamber using a high-resolution digital imaging system. Richard Alexander, from Stowers microscopy center helped quantitate the colony sizes by measuring the area of each colony (in pixels squared) using an Axiovision script. A computational pipeline for processing SGA raw data was then created to identify quantitative genetic interactions by applying a series a normalization to correct for numerous systematic experimental effects like Plate-specific effect, Row/column effect, Spatial effect, competition effect etc. At least 2 of the four quadruplicate spots should have grown in the control plate for that gene to be considered. For each spot, a Z score was calculated using the equation (x- mean of all samples)/ (standard deviation of all samples), where x is the intensity of a given spot. For a spot to be considered significantly down, z score should be less than 2*(- standard deviation) with a median change of <0.5 for atleast 2 of the spots. A list of significantly sick genes or the negative genetic interactions for each experiment was created.

5.3 Interpretation and analysis of negative genetic interactions

We performed the screen in triplicate for all three deletion mutants. From these screens, we found two classes of genes; (1) show negative genetic interactions with all three genes *ell1*, *eaf1* and *ebp1*, and (2) show strong genetic interaction with *ell1*, and very little with the *eaf1* or *ebp1*, again raising the possibility that Ell1 may exert some

functions independent of the other factors. The latter observation is perhaps not surprising since, since Ell1 binds directly to Pol II and at least *in vitro* can weakly enhance elongation on its own. Some of the interesting hits from both the classes are listed below. Proteins that act separately but have overlapping important functions show negative genetic interaction or synthetic lethality, such that loss of both genes results in severe growth defect. The components do not act in the same pathway.

1) Genes interacting genetically with all three strains; *ell1Δ*, *eaf1Δ* and *ebp1Δ*:

A group of genes show negative genetic interactions in many of our screens with the three genes tested. Table 5.1 shows some of these genes. The list is extensive. We found one of the most interesting hits to be *brl1*. *brl1* I encodes a ubiquitin protein ligase was found in 6 of the 9 screens, genetically interacting with *ell1*, *eaf1* and *ebp1*. We chose to pursue it because of its obvious connection to transcription and chromatin structure. Brl1 is essential for histone H2B ubiquitylation and is known to be important for recruitment of P-TEFb kinase (Sanso et al., 2012a). I will discuss the interaction between *brl1Δ* and *ell1Δ*, *eaf1Δ*, and *ebp1Δ* in the next section (Section 5.4).

Table 5.1 Negative genetic interaction

SGA screens for *ell1Δ*, *eaf1Δ* and *ebp1Δ* were performed in triplicate. Some of the hits for each screen are shown here. '0' represents no detectable genetic interaction whereas a positive hit or a 'negative genetic interaction' for a given screen is represented by '1' (red box).

gene	product	Ell1			Eaf1			AF4-L		
ogm1	protein O-mannosyltransferase Ogm1	1	1	0	1	1	1	1	1	1
brl1	ubiquitin-protein ligase E3 Brl1	1	1	1	0	1	1	0	1	0
rps1101	40S ribosomal protein S11 (predicted)	1	1	1	0	0	0	1	0	1
atg24	autophagy associated protein Atg24 (predicted)	1	1	0	1	1	0	0	1	1
efr3	HEAT repeat protein involved in protein localization to plasma membrane	1	1	1	0	1	1	0	1	1
rav1	RAVE complex subunit Rav1	1	1	0	1	1	1	0	1	0
new4	conserved eukaryotic protein	1	1	1	0	0	1	0	1	0
rpn15	proteasome regulatory particle, lid subcomplex subunit Rpn15/Dss1	1	1	1	0	1	0	0	0	1
alm1	medial ring protein Alm1	1	0	0	0	0	1	1	1	1
ace2	transcription factor Ace2	1	1	0	0	0	0	1	1	1
kin1	microtubule affinity-regulating kinase Kin1	1	1	1	0	1	1	0	1	1
sim3	NASP family CENP-A chaperone	1	1	1	0	1	0	0	1	0
cap1	adenylyl cyclase-associated protein Cap1	1	1	1	0	1	1	0	0	1
lsm1	mRNA decapping complex subunit (predicted)	1	1	1	0	1	0	0	0	1
fyv7	rRNA processing protein Fyv7 (predicted)	1	1	1	0	0	0	0	0	1
nup60	nucleoporin Nup60	1	1	0	1	1	1	0	0	0
pus7	pseudouridine synthase Pus7 (predicted)	1	1	1	1	0	0	0	0	0
arg4	arginine specific carbamoyl-phosphate synthase Arg4 (predicted)	1	1	1	0	1	1	0	0	0
cnb1	calcineurin regulatory subunit (calcineurin B)	1	1	1	0	1	0	0	0	0
dbp7	ATP-dependent RNA helicase Dbp7 (predicted)	1	1	1	0	0	1	0	0	0
SPAPB1E7.11c	Schizosaccharomyces specific protein	1	1	0	0	0	0	1	1	1
rpl3202	60S ribosomal protein L32 (predicted)	1	1	0	0	0	0	1	1	0
SPBC16C6.01c	lysine methyltransferase (predicted)	0	0	1	0	0	0	1	0	1
png1	ING family homolog Png1	1	1	0	0	0	0	0	1	1
ect1	ethanolamine-phosphate cytidylyltransferase (predicted)	1	1	0	0	0	0	0	1	1
clu1	clustered mitochondria (cluA/CLU1) homolog Clu1 (predicted)	1	1	1	0	1	1	0	1	0
rpl2002	60S ribosomal protein L20 (predicted)	1	1	0	0	1	0	0	1	0
prm1	conjugation protein Prm1	1	1	0	0	0	1	0	1	0
rpl1002	60S ribosomal protein L10	1	1	0	0	0	1	0	1	0
SPBC13G1.14c	RNA-binding protein (predicted)	1	1	1	0	0	0	0	1	0

2) Genes that genetically interact with only *ell1Δ*:

The screens identified a large number of genes that interact genetically with *ell1Δ* but not with *eaf1Δ* or *ebp1Δ*, raising the possibility that Ell1 has functions that are partially or wholly independent of its binding partners Eaf1 and/or Ebp1. This observation is somewhat surprising in light of the finding that deletion of eaf1 decreases ell1 expression levels; however, it suggests that enough residual ell1 protein remains in the eaf1 deletion strain to support key ell1 functions. As noted above, we know Ell1 is the only one of the

three that can directly bind Pol II, and *in vitro* experiments show that Ell1 can stimulate rate of elongation by itself. Hence, a larger list of genetic interaction may be an indication that the Ell1 protein is more important for cellular functions than Eaf1 or Ebp1. One interesting group of genes that fell into this category were multiple genes involved in heterochromatin formation and maintenance. These interactions are explored further later in the chapter.

5.4 Genetic Interaction between *brl1Δ* and *ell1Δ*, *eaf1Δ*, and *ebp1Δ*

The *brl1* gene was among the genes found to have negative genetic interaction with *ell1*, *eaf1* and *ebp1* in the SGA screens. Brl1 was of particular interest as an SGA hit as it has genetic interactions with genes encoding all three of the ELL complex components, and because Brl1 has role in maintaining the chromatin state and in transcription in cells. Brl1 is a ubiquitin protein ligase E3 that puts a mono-ubiquitin mark at the K119 position of histone H2B. Mono-ubiquitylation of histone H2B (H2Bub1) is required for co-transcriptional generation of H3K4 marks by the methyltransferase Set1 in yeast, and it contributes to global H3K4me levels in metazoans. H2Bub1 also regulates gene expression through an unidentified, methylation-independent mechanism. Notably, H2B ubiquitylation has been linked to Pol II elongation. For example, in metazoa, cotranscriptional monoubiquitylation of H2B and the Pol II elongation rate are tightly coupled (Fuchs et al., 2014). In *S. pombe*, H2Bub1 and Cdk9 (P-TEFb) are regulated by a positive feedback loop, where Cdk9 activity is needed for co-transcriptional H2B ubiquitylation, and H2Bub1 in turn stimulates Cdk9 activity toward Spt5 (Sanso et al.,

2012b). H2Bub1 is also needed for centromeric chromatin maintenance, promoting noncoding transcription, centromere integrity, and accurate chromosomal segregation (Sadeghi et al., 2014a). It is also interesting to note that multiple proteins involved in centromeric integrity during chromosome segregation also showed up in the SGA screens (Table 5.2). This further suggested that Ell complex and Brl1 functions might be linked.

gene	function	*ell1Δ*			*eaf1Δ*			*ebp1Δ*		
kin1	microtubule affinity-regulating kinase Kin1	1	1	1	0	1	1	0	1	1
sim3	NASP family CENP-A chaperone	1	1	1	0	1	0	0	1	0
mhf2	CENP-X homolog, FANCM-MHF complex subunit	1	1	1	0	0	0	0	0	0
cnp3	CENP-C ortholog Cnp3	0	0	1	0	0	1	0	0	0
spc19	DASH complex subunit Spc19	0	0	0	1	1	0	0	0	0
duo1	DASH complex subunit Duo1	1	1	0	0	0	0	0	0	0

Table 5.2 *ell1*, *eaf1* and *ebp1* interact genetically with multiple genes implicated in centromeric integrity

SGA screen for the *ell1Δ*, *eaf1Δ* and *ebp1Δ* was performed in triplicates and these proteins were seen to interact with multiple genes important for the centromeric architecture

To further confirm the synthetic sickness of the Ell complex-*brl1* double mutants, we made the following strains independently; *ell1Δ brl1Δ*, *eaf1Δ brl1Δ* and *ebp1Δ brl1Δ*. We checked the growth of these strains in liquid medium using TECAN microplate reader, which allows measurement of the absorbance of the culture with time. Single deletions of *ell1+*, *eaf1+* and *ebp1+* had no effect on cell growth in rich medium, while cells carrying the *brl1+* deletion grew more slowly than wild type. As shown in Figure 5.2, the double deletion strains grew much more slowly than wildtype or the single *brl1Δ* strains, again suggesting that the ELL complex has strong genetic interaction with *brl1*.

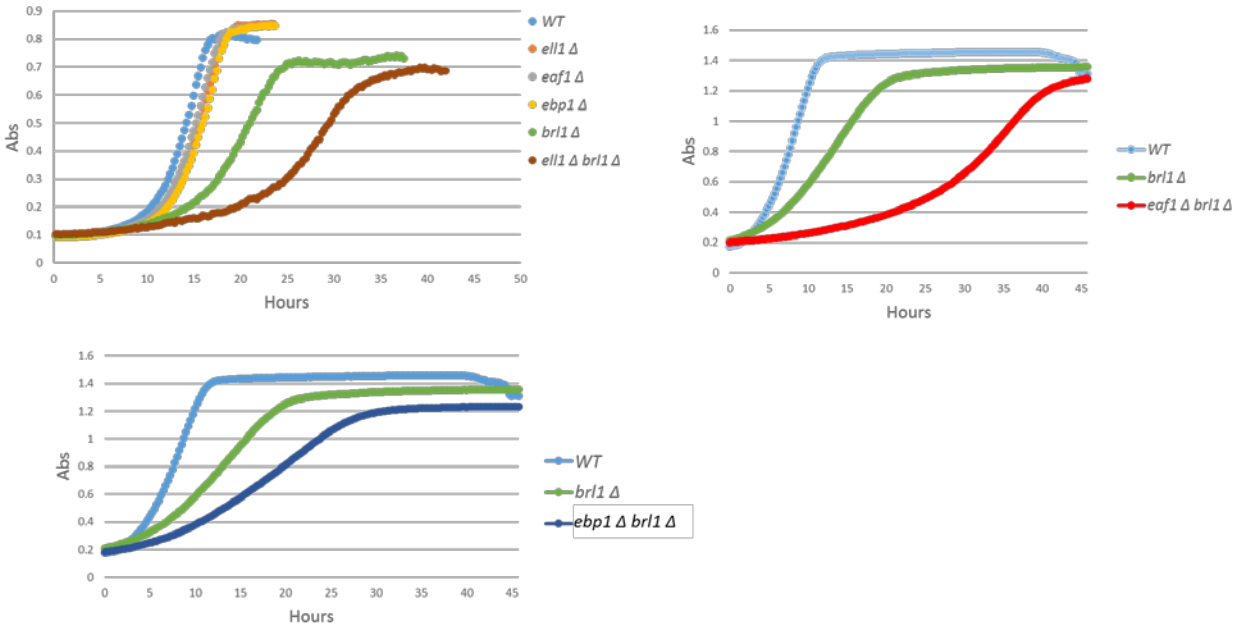


Figure 5.2 *ell1+*, *eaf1+* and *ebp1+* deletion when combined with *brl1+* deletion makes the cells sick.

Graphs showing change in absorbance with time for various strains tested. The slope of the graph denotes the growth rate.

5.5 Effect of *ell1+*, *eaf1+* and *ebp1+* deletions on Brl1 function

Brl1 plays a key role in H2B ubiquitylation and, because it helps to recruit Set1, the subsequent Set1-dependent H3K4 methylation. For these reasons, we wanted to determine whether deletion of *ell1+*, *eaf1+* and *ebp1+* would have any effect on overall H2B mono-ubiquitylation or H3K4 methylation in cells. For this, we made whole cell extracts from all strains and performed western analysis using antibodies specific for H2Bub1 and H3K4me1. As shown in Figure 5.3, I found that deletion of *ell1+*, *eaf1+* or *ebp1+* had no effect on the overall levels of H2Bub1 and H3K4me1. A strain in which H2B cannot be ubiquitylated by Brl1 was used as a control to ensure that we observe the expected loss of H3K4me1 and H2Bub1 in these assays.

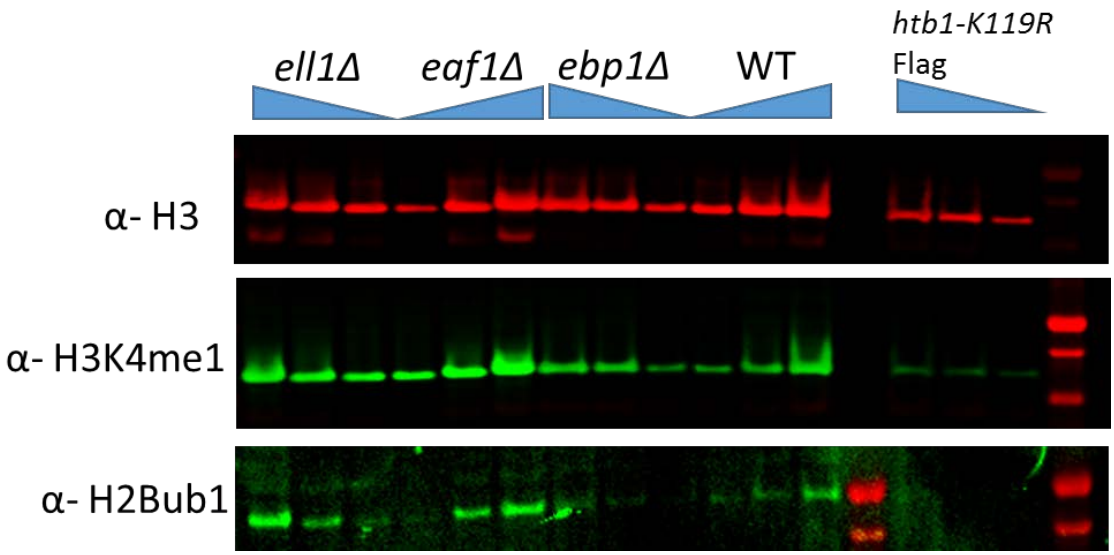


Figure 5.3 *ell1+*, *eaf1+* or *ebp1+* deletion does not affect the overall levels of H3K4me1 or H2Bub1

Western analysis using specific antibodies against H3 (used as loading control), H3K4me1 and H2Bub1. Different amounts of whole cell extracts for each strain were loaded and the blots were probed with indicated antibody.

Because H2B is the major target for Brl1 ubiquitin ligase, we anticipated that double mutant strains carrying the *htb1-K119R* mutation along with the *ell1+*, *eaf1+* or *ebp1+* deletions would behave similarly the Brl1-ELL complex double mutant strains. To our surprise, however, we found that deletion of the ELL complex in the *htb1-K119R* mutant strain had no effect on the growth of cells (Figure 5.4), as the double mutants grow similarly to the *htb1-K119R* single mutant. This observation suggests that the genetic interaction we observe between the ELL complex and Brl1 could be due to ubiquitylation-independent mechanisms.

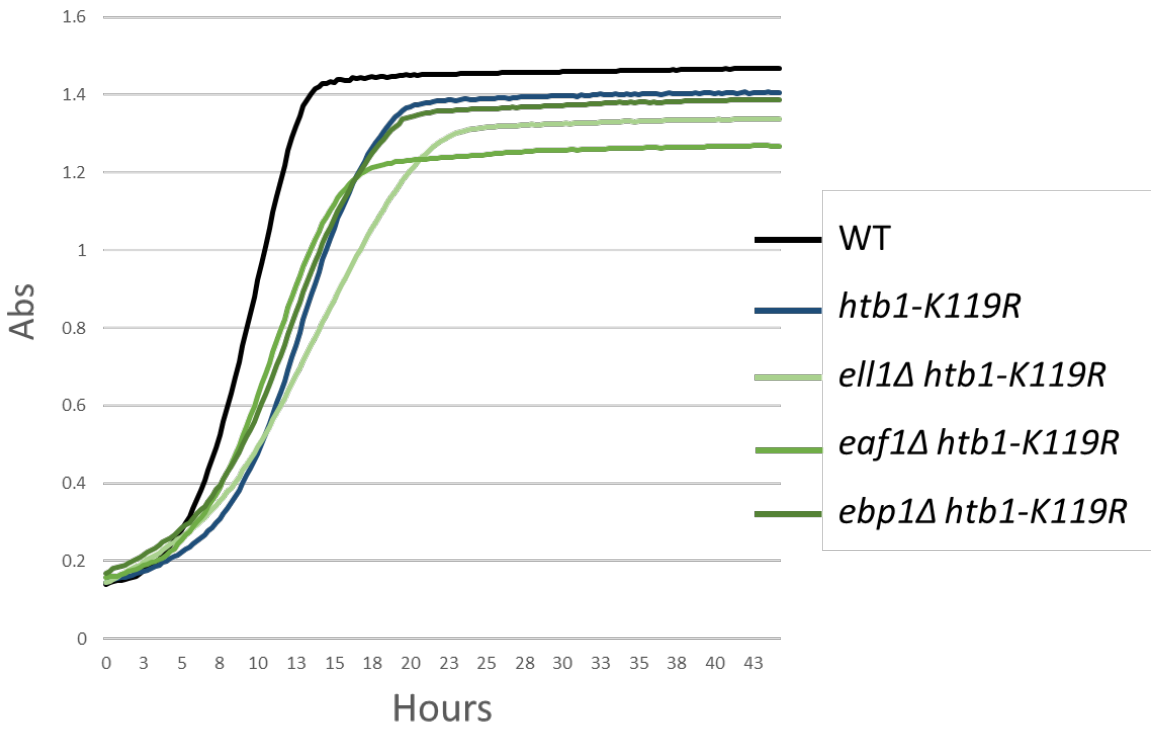


Figure 5.4 *ell1+*, *eaf1+* and *ebp1+* deletion combined with *htb1-K119R* mutation does not mimic the growth defects seen with brl1 deletion

Graphs showing change in absorbance with time for various strains tested. The slope of the graph denotes the growth rate.

Previous experiments have suggested a role of H2Bub1 in mitotic progression by using *htb1* mutants with affected ubiquitylation status. We therefore wished to determine whether mutants in ELL complex components exhibit mitotic phenotypes. First, we tested whether deletion of *ell1+* affects centromere function by growing mutant strains in the presence of the microtubule-destabilizing drug thiabendazole (TBZ). Figure 5.5 shows that as previously reported (Sadeghi et al., 2014b) H2Bub1-deficient cells (*htb1-K119R*) were sensitive to TBZ, thus confirming that H2B monoubiquitylation acts synergistically with TBZ during chromosome segregation. Of note, deletion of *ell1+* also rendered cells sensitive to TBZ.

Mutations causing loss of H2Bub1 are also known to increase chromosomal mis-segregation. For a direct measurement of genome stability, we sought to analyze the chromosome segregation in M phase cytologically. We DAPI stained the cells to look for aberrant chromosomal segregation. As expected, the *brl1Δ* and *htb1-K119R* exhibited increased (OR severe) chromosomal lagging and stretching leading to mis-segregation; however, the *ell1Δ*, *eaf1Δ* and the *ebp1Δ* strains behaved similar to the wildtype strain (data not shown).

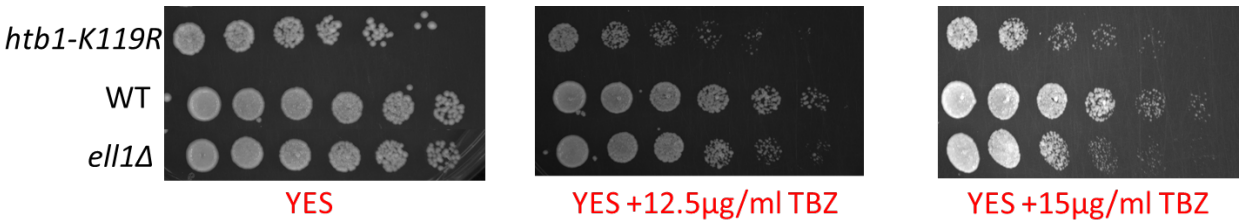


Figure 5.5 TBZ sensitivity of *ell1Δ* strain.

The Wildtype, *ell1Δ* and *htb1-K119R* strains, were grown to mid log phase in rich media, washed in 1X PBS, and resuspended in H_2O at a density of 1 x 10^8 cells/ml. 5 µl of 3-fold serial dilutions of cells were spotted onto YES plates with or without thiabendazole.

Deletion of *brl1+* affects recruitment of Cdk9 kinase (Sansó et al., 2012). As I have shown, Cdk9 and the ELL complex co-occupy many genomic loci, prompting us to determine whether Brl1 or H2B ubiquitylation are needed for optimal ELL complex recruitment in *S. pombe*. We performed ChIP-qPCR analysis to test whether recruitment of ELL complex is affected in *brl1Δ* and *htb1-K119R* strains, and found that absence of H2Bub1 had no effect on the recruitment of Ell1, Eaf1 or EBP1 (data not shown).

5.6 *ell1Δ* shows negative interaction with genes involved in heterochromatin formation

As discussed in Chapter 5, I found that deletion of *ell1+* led to a significant decrease in H3K9me2 at subtelomeric regions of chromosomes 1 and 2. In addition, I detected an increase in expression of a number of genes within the affected loci in the *ell1Δ* strain, suggesting that Ell1 could play a role in heterochromatin formation or maintenance. Heterochromatin formation is mediated by distinct histone-modifying enzymes (H3K9me), RNAi proteins, components of the Pol II machinery and homologs of heterochromatin protein 1 (HP1) (Allshire and Ekwall, 2015). These mechanisms are conserved in higher eukaryotes.

A major hallmark of heterochromatin in most eukaryotes is the methylation of histone H3 at lysine 9 (H3K9) and histone hypocetylation. In fission yeast, H3K9 methylation is carried out by a histone methyltransferase, Clr4, which is responsible for mono-, di- and tri-methylation of H3K9. H3K9me forms a binding site for chromodomain-containing heterochromatin proteins, including the HP1 homologs Swi6 and Chp1, and importantly, for Clr4 itself, leading to models for perpetuation and spreading of heterochromatin (Ekwall et al., 1996; Ivanova et al., 1998; Zhang et al., 2008). Clr4 can be isolated as part of a complex containing a cullin 4 (Cul4)-based E3 ubiquitin ligase and Rik1 protein, termed the CLRC (CLr4-Rik1-Cul4) complex, that help in heterochromatin formation (Hong et al., 2005). Heterochromatic silencing also requires multiple conserved deacetylases (HDACs) like Clr3, Clr6 and Sir2, and the components of the RNAi machinery. Among these are the Argonaute-containing RNA-induced transcriptional silencing (RITS) RNAi effector complexes (Sadaie et al., 2004). RITS is required for H3K9 methylation by

Clr4 and conversely, H3K9 methylation can attract RITS to chromatin via binding of the Chp1 protein. RITS also interact with RNA-directed RNA polymerase (RDRC) whose activity is critical for siRNA generation and heterochromatin assembly (Motamedi et al., 2004).

Interestingly, when we performed SGA screen with *ell1Δ* strain, we found that a collection of genes encoding proteins involved in heterochromatin formation or maintenance were identified as negative interactors with *ell1Δ*. These interactions did not seem to occur with *eaf1Δ* or *ebp1Δ*. This finding is particularly interesting in light of our observation that heterochromatin formation is altered specifically in *ell1Δ* cells and is consistent with the idea that Ell1 may have a role independent of Eaf1 and Ebp1 in maintaining heterochromatin in S*. pombe*. Table 5.3 show some of these interactions.

gene	function	*ell1Δ*			*eaf1Δ*			*ebp1Δ*		
ago1	Argonaute	1	1	1	0	0	0	0	0	0
laf2	Clr6 associated factor	1	1	1	0	0	0	0	0	0
rik1	Silencing protein	1	1	1	0	0	0	0	0	0
clr4	Histone H3 methyltransferase	1	1	0	0	0	1	0	0	0
clr1	SHREC complex subunit	1	0	1	0	0	0	0	0	0
clr3	Histone deacetylase (class II)	0	1	1	0	0	0	0	0	0

Table 5.3 *ell1Δ* interacts genetically with genes implicated in heterochromatin formation and/or maintenance

SGA screens with *ell1Δ*, *eaf1Δ* or *ebp1Δ* strains were performed in triplicates. Only *ell1Δ* interacted reproducibly with multiple genes involved in heterochromatin formation and/or maintenance.

5.7 *ell1Δ* interactions in presence of mycophenolic acid

From work presented in this thesis and elsewhere (Banks et al., 2007), we know that strains deleted for *ell1* exhibit sensitivity to drugs such as mycophenolic acid and 6-azauracil, phenotypes associated with a number of transcription elongation regulators. To determine whether *ell1Δ* would interact with additional genes when grown under conditions thought to compromise transcription elongation, we performed additional SGA screens on MPA-containing medium.

We performed the SGA screen in triplicates as before with *ell1Δ* cells, but at the last plating used plates containing 15 µg/ml mycophenolic acid in addition to the selection markers. We identified a few genes that interacted with *ell1Δ* only in the presence of MPA and that could possible play a role in transcription (Table 5.4). Ssr4 is a part of Chromatin remodelers like SWI/SNF and RSC and can hence modulate transcription. Cdk11 phosphorylates the Med27 and Med4 Mediator subunits on conserved residues and is required for the assembly of Mediator complex in *S. pombe* (Drogat et al., 2012). Studies indicate that Rhp41, a fission yeast homolog of human XPC, is involved in transcription-coupled repair of MMS-induced DNA damage (Kanamitsu and Ikeda, 2011). Splicing associated factor, Saf4, is an interesting hit because it splicing can influence the rate of transcription elongation. We also got hits like the protein coded by SPAC18B11.11 that have no known transcription role, but showed up in all our screens in the presence of MPA. In the future, it would be worth analyzing these interactions further as they may shed light on elongation specific functions of Ell1.

gene	Product	*ell1Δ*			*ell1Δ* + MPA		
ssr4	SWI/SNF and RSC complex subunit Ssr4	0	0	0	1	1	1
SPAC18B11.11	GTPase activating protein (predicted)	0	0	0	1	1	1
cdk11	serine/threonine protein kinase cdk11	0	0	0	1	1	0
rhp41	DNA repair protein Rhp41	0	0	0	1	1	0
saf4	splicing associated factor Saf4	0	0	0	1	1	0

Table 5.4 *ell1Δ* interacts genetically with some genes only in the presence of MPA

SGA screen for the *ell1Δ* was performed in triplicates with or without MPA (15 ug/ml) and some interactions were seen only in the presence MPA.

5.8 Discussion

A genetic interaction occurs when the combination of two mutations leads to an unexpected phenotype. Screens for synthetic genetic interactions have been used extensively to identify genes whose products are functionally related (Baryshnikova et al., 2013) Using a method termed synthetic genetic array, is a high-throughput technique for identifying genetic interactions (Baryshnikova et al., 2010), we performed three separate screens – each in triplicate – to identify genes that interact with the *ell1*, *eaf1*, and *ebp1* genes. SGA is a powerful technique for large-scale construction of mutants and assessment of phenotypic consequences associated with combinatorial genetic perturbations. From our screens, we obtained a list of genes that show negative genetic interactions with *ell1Δ*, *eaf1Δ* or *ebp1Δ* strains. Many of the genes identified were identified in all three screens, and some were involved in related pathways We focused first on the Brl1 gene, which encodes an E3 ligase involved in H2B ubiquitination at K119 position in *S. pombe*.

H2Bub1 levels in chromatin generally correlate with Pol II and transcription but H2Bub1 has also been implicated in diverse cellular functions, such as DNA replication, differentiation and DNA-damage responses. H2B monoubiquitylation, which appears to play a central role in the interplay between chromatin and the RNAPII elongation complex, is catalyzed in budding yeast by the ubiquitin-conjugating enzyme Rad6 and the E3 ubiquitin ligase Bre1 (Osley, 2004). Formation of H2Bub1 on transcribed chromatin also requires PAFc (Jaehning, 2010). H2Bub1 is important for co-transcriptional generation of H3K4me by the methyltransferase Set1 in yeast (Racine et al., 2012; Sun et al., 2015; Sun and Allis, 2002), and contributes to global H3K4me levels in metazoans (Dover et al., 2002). A positive feedback loop, in which phosphorylation by Cdk9 stimulates H2Bub1, and H2Bub1 stimulates chromatin association of Cdk9 and phosphorylation of Spt5 (DSIF subunit), is reported to regulate elongation in *S. pombe* (Sanso et al., 2012b). Noncoding centromeric transcription is dependent on H2Bub1, and in H2Bub1-deficient cells, centromere functionality is hampered resulting in unequal chromosome segregation (Sadeghi et al., 2014b).

We generated independent double mutants having deletion of *brl1+* with *ell1Δ*, *eaf1Δ* or *ebp1Δ* and confirmed that the double mutants indeed grow much slower than the wildtype or the single mutants. But surprisingly, lack of H2Bub1 in a mutant that cannot monoubiquitinate H2B did not incite the same growth phonotype when combined with *ell1Δ*, *eaf1Δ* or *ebp1Δ*. This suggests that the genetic interaction that *ell1Δ*, *eaf1Δ* or *ebp1Δ* show with *brl1Δ* could be due to functions independent of H2B ubiquitination. Furthermore, we found that the overall levels of H2b ubiquitination or H3K4 methylation were not detectably changed upon *ell1+*, *eaf1+* or *ebp1+* deletion; whether the ELL complex contributes to placement of these marks at specific locations within the genome

remains to be determined. Deletion of *ell1+* did show a sensitivity when grown in presence of the microtubule-destabilizing drug thiabendazole, but what if any role *ell1+* may play in chromosome segregation needs to be tested further.

ell1Δ showed genetic interaction with many more genes than *eaf1Δ* or *ebp1Δ*, which may suggest that Ell1 has functions independent of Eaf1 and Ebp1 and/or plays a more central role in processes regulated by the ELL complex than Eaf1 and Ebp1. Consistent with this observation, only deletion of *ell1+* causes a sensitivity to 6-azauracil and mycophenolic acid, and studies show that Ell1 by itself can stimulate *in vitro* elongation by Pol II, suggesting that it may have a bigger role than Eaf1 and Ebp1. Some of the genes/pathways that showed up as genetic interactors of *ell1Δ* were interesting and worth studying further in the future. In further experiments, we performed the SGA screen for *ell1Δ* in the presence of MPA and found identified additional interactions detected only under these conditions. Some of the interactions that we found only in the presence of MPA are SWI/SNF and RSC subunit Ssr4, DNA repair protein Rhp41 and splicing associated factor Saf4. These genetic interactions can be studied more in the future as they may help to elucidate the role of Ell1 in transcription elongation.

It was interesting to find that that multiple genes with roles in heterochromatin formation or maintenance were identified in our SGA screens. Argonaute or Ago1 in *S. pombe* showed up in all 3 *ell1Δ* SGA screen. Ago1, a component of RNA-induced transcriptional silencing complex (RITS), is needed for co-transcriptional gene silencing by RNA interference machinery. RITS is proposed to be the key initial trigger for heterochromatin assembly and regulation of histone H3K9 methylation. Clr4 (Cryptic loci regulator 4) is the only histone H3K9 methyltransferase in *S. pombe* and therefore is

absolutely required for heterochromatin regulation. Clr4 has been found to be part of a multiprotein complex termed “Clr4 methyltransferase complex” (CLRC). CLRC is composed of Clr4, Raf1, Raf2, Cul4, and Rik1. Rik1, Clr1, Clr2, Clr3, Clr4 and Swi6 were all identified in S. pombe screens for factors involved in heterochromatin silencing (Ekwall and Ruusala, 1994; Thon and Klar, 1992). Clr1, Clr2 and Clr3 are part of the Snf2/HDAC-containing Repressor Complex (SHREC) that mediates heterochromatin transcriptional gene silencing in fission yeast (Sugiyama et al., 2007). Clr3 is a an H3-K14-specific histone deacetylase II. SHREC is recruited to the telomeres by multiple independent mechanisms and acts to restrict Pol II occupancy at heterochromatin repeats. SHREC interacts with Ccq1, a telomere binding protein that along with Taz1 acts in a parallel mechanism to RNAi pathway to facilitate recruitment of SHREC to the telomeric ends (Moser et al., 2015). Clr4, Rik1, Clr1 and Clr3 were all found to be strong negative genetic interaction in the ell1Δ screens, again consistent with the possibility that Ell1 plays a role in heterochromatin maintenance. In the future, it will be of considerable interest to explore the connections between Ell1 and the genes and proteins regulating heterochromatin formation and maintenance.

CHAPTER 6: Conclusion

This thesis elucidates the role of Ell1 and Eaf1 in *S. pombe* that act as elongation factors and share functional similarities with the ELL-EAF complex found in higher eukaryotes. Since *S. pombe* has genes encoding P-TEFb, Ell1 and Eaf1 we hypothesized that a larger elongation complex like the Super elongation complex in mammals might be present. Proteomic experiments found that both Ell1 and Eaf1 interact with a previously uncharacterized protein encoded by the gene SPAC6G9.15c. Upon performing Psi-blast with relaxed stringency we find that that this new protein, that we refer to as Ell1 binding protein (Ebp1) might have sequences of faint homology to the AF4 family of proteins from higher eukaryotes. We showed that Ell1, Eaf1 and Ebp1 can stably interact and are recruited together to places having high Pol II and P-TEFb (cdk9), suggesting that fission yeast may have an SEC-like complex. In-vitro transcription assays performed using recombinant proteins showed that while Ell1/Eaf1 can stimulate the rate of elongation as previously reported, the addition of Ebp1 had no detectable effect on the elongation by Pol II under the conditions tested, indicating that Ebp1 plays a role undetectable in our assays. Ebp1 might play a potential role as a scaffold on which other proteins can assemble, or by helping to recognize chromatin marks, or by linking Ell1/Eaf1 to other regulatory proteins. Deletion of Ell1 causes a 5' shift in the pattern of Pol II occupancy across genes, which is consistent with the idea that it acts as an elongation factor such that the deletion of Ell1 causes Pol II to lag behind. In the future, it would be interesting to check whether the loss of Ell complex might lead to changes in CTD S2 phosphorylation, a mark that is placed by the P-TEFb kinase Cdk9 and is diminished in higher eukaryotes upon ELL depletion. The deletion of these genes had minimal effect on

steady state levels of messenger and other RNAs in cells grown in rich media under optimal conditions, though it is worth noting that a group of genes regulated by the transcription factor Ace2, including the Ace2 transcript itself were downregulated. These Ace2 and Ace2 regulated genes also exhibit disproportionally high Ell complex binding relative to Pol II compared to the genome-wide average, suggesting that Ell complex could play a direct role in the regulation of expression of Ace2 and Ace2 regulated genes. It would be interesting to check whether the absence of Ell1, Eaf1 or Ebp1 influences the cell cycle stages or causes septation defects as in *ace2Δ* cells.

Genetic screens performed to identify the proteins that may work in similar pathways (negative genetic interactions) gave us multiple targets that interacted genetically with either all three of the proteins or Ell1 alone. It is perhaps not surprising that deletion of Ell1 has a more prominent phenotype, as Ell1 is the subunit that can directly bind to Pol II and stimulate the rate of transcription elongation by itself.

It was interesting to find that that multiple genes with roles in heterochromatin formation or maintenance were identified in our SGA screens, suggesting that Ell1 may play a role in heterochromatin maintenance. Heterochromatin distribution was checked by H3k9me2 ChIP, and it was found that deletion of Ell1 gene changes the pattern of heterochromatin at the sub -telomeric regions of the chromosomes 1 and 2. Interestingly, a few CTD mutants have also been reported to give rise to similar H3K9methylation phenotypes as ell1 deletion (Inada et al., 2016). It is interesting to note that in RNA seq a cluster of genes in the region close to the right arm of Chromosome 1 (where the H3K9 methylation mark is reduced) seemed to be modestly upregulated upon *ell1+* deletion, again reinforcing the possibility that Ell1 may play a role in

heterochromatin regulation/maintenance at the sub-telomeric regions. Reintroducing ell1+ into the mutant strains to see whether the phenotype can be rescued might validate the role of Ell1 in heterochromatin formation and maintenance.

CHAPTER 7: Materials and Methods

7.1 *S. pombe* strain construction

The gene manipulations necessary to generate fission yeast strains required for mass spectrometry and EM analyses were carried out using standard protocols. The strains used in this study are listed in Appendix A. A PCR-based genomic epitope-tagging method was used to construct subunit deletion strains, where coding region of the gene of interest was replaced by a selection marker (NatMX6 or KanMX6).

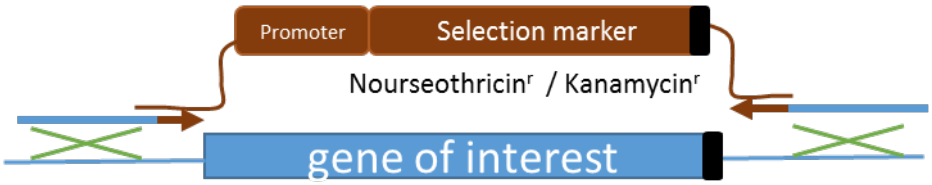


Figure 7.1 Construction of gene deletion strains.

PCR primers (indicated by arrows) were used with the pFA6a plasmid as template to generate a PCR product with 90bp of homology to regions flanking the sequence to be deleted as indicated. Transformation of the PCR product into S.pombe cells resulted in replacement of the target sequence with a Nourseothricin or kanamycin resistance cassette under the control of promoter sequences from Ashbya gossypii (black box represents the STOP codon). Transformed cells were plated and isolated on YES containing 200 mg/l clonNAT/ G418. Deletion of the target sequence and insertion of the resistance cassete at the correct locus were both confirmed by PCR.

Strains used for MudPIT analysis were C-terminally Flag-tagged on by replacing the stop codons of each gene with sequences encoding 3× Flag followed by a stop codon and the NatMX6 marker (Bahler et al., 1998).

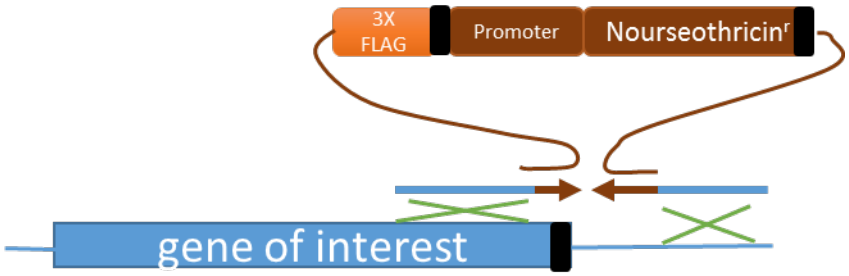


Figure 7.2 Construction of epitope tagged strains.

The target gene is indicated as the blue box with the stop codon (black). PCR primers were designed with 90 bases of homology to the region immediately upstream of the stop codon (5') and a region downstream of the stop codon (3'). These were used to generate a PCR product using the pFA6a 3X FLAG natMX6 plasmid as a template. This includes 3 tandem repeats coding for the flag epitope with a transcription termination sequence followed by a nourseothricin antibiotic cassette. The PCR product was used to generate recombinant strains as described in Figure 7. .

7.2 Immunopurification for Mass spectrometry

S. pombe Mediator and Ell1/Eaf1 preparations used for MudPIT mass spectrometry were purified by anti-Flag immunopurification from 972 h- derivatives expressing Flag-tagged Mediator subunits. Flag-immunopurified proteins were treated with benzonase, TCA precipitated, denatured, reduced and alkylated, digested with endoproteinase Lys-C and trypsin, and analyzed by MudPIT mass spectrometry.

972 h- cells or 972 h- derivatives expressing Flag-tagged proteins were grown at 32 °C in rich medium (YES) supplemented with adenine, histidine, leucine and uracil (225 µg/ml). Cells were harvested at mid-log phase, washed in cold H_2O, and washed in buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% NP40, and 10% glycerol with protease inhibitor cocktail (Sigma P8849). Cells were pulverized in liquid nitrogen using a mortar and pestle and then re-suspended in the same buffer. Whole-cell lysates were centrifuged at 40,000 rpm to remove cell debris. Supernatants were incubated overnight

at 4°C with anti-Flag agarose beads (Sigma, 500µl beads per 10^8 starting cells) that had been pre-equilibrated in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% NP40, and 10% glycerol with protease inhibitor cocktail. Beads were washed four times with approximately 10 bead volumes of buffer containing 50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 10% glycerol and 0.1% NP40, and bound proteins were eluted using 0.2 mg/ml 3× Flag peptide (Sigma) in 50 mM Tris-HCl, 300 mM NaCl, 10% glycerol and 0.1% NP40. Flag-immunopurified proteins were treated with benzonase and TCA precipitated before analysis by multidimensional protein identification technology (MudPIT).

TCA-precipitated proteins were urea-denatured, reduced, alkylated and digested with endoproteinase Lys-C (Roche) followed by modified trypsin (Roche) (Washburn et al., 2001). Peptide mixtures were loaded onto 100-µ m fused silica microcapillary columns packed with 5-µm C18 reverse phase (Aqua, Phenomenex), strong cation exchange particles (Partisphere SCX, Whatman), and reverse phase (Florens and Washburn, 2006). Loaded microcapillary columns were placed in-line with an LCQ or LTQ ion trap mass spectrometer equipped with a nano-LC electrospray ionization source (ThermoFinnigan). Fully automated MudPIT runs were carried out on the electrosprayed peptides. Tandem mass (MS/MS) spectra were interpreted using SEQUEST against a database of a database of S. pombe proteins (downloaded from NCBI on 23 January 2012), and complemented with 177 sequences from usual contaminants (human keratins, IgGs, proteolytic enzymes). To estimate false positive discovery rates, each sequence was randomized keeping amino acid composition and length the same, and the resulting ‘shuffled’ sequences were added to the ‘normal’ database (doubling its size) and searched at the same time. Peptide/spectrum matches were sorted and selected using DTASelect with the following criteria set: spectra/peptide matches were only

retained if they had a DeltaCn of at least 0.08, and minimum XCorr of 1.8 for singly, 2.0 for doubly, and 3.0 for triply charged spectra. In addition, peptides had to be fully tryptic and at least 7 amino acids long. Combining all runs, proteins had to be detected by at least 2 such peptides or by 1 peptide with 2 independent spectra. Under these criteria, the overall false discovery rate was less than 0.2%. Peptide hits from multiple runs were compared using CONTRAST (Tabb et al., 2002). To estimate relative protein levels, distributed normalized spectral abundance factors (dNSAFs) were calculated for each detected protein (Florens et al., 2006).

7.3 Expression and purification of Recombinant Proteins in Insect Cells

cDNAs encoding wild-type *ell1+*, *eaf1+* and *ebp1+* containing N-terminal FLAG or c-Myc epitope tags were subcloned into pBacPAK8. Recombinant baculoviruses were generated with the BacPAK expression system (Clontech) using pBacPAK6 viral DNA prepared as described (Kitts and Possee, 1993), and stored in 50ml centrifuge tubes spray-painted black to prevent exposure to light during storage. Sf21 insect cells were cultured at 27 °C in Sf-900 II SFM (Invitrogen). Flasks containing 1 x 10^8 Sf9 cells were infected with the recombinant baculoviruses. Forty-eight hours after infection, cells were collected and lysed in 15 ml of ice-cold buffer containing 50 mM Hepes-NaOH (pH7.9), 300mM NaCl, 5 mM $MgCl_2$, 0.2% Triton-X-100 and 20% (v/v) glycerol with protease inhibitor cocktail. Lysates were centrifuged 40,000 rpm for 30 min at 4°C.

FLAG-tagged proteins were purified from Sf9 cell lysates by anti-FLAG agarose immunoaffinity chromatography. Lysates from 1 x 10^8 cells were incubated with 0.5 ml anti-FLAG (M2) agarose beads overnight at 4°C. The beads were washed three times with Tris-buffered saline (TBS, 25 mM Tris-HCl (pH 7.4), 137 mM NaCl, 2.7 mM KCl), and bound proteins were eluted from the beads with TBS containing 10% (v/v) glycerol and 0.3 mg/ml 3X FLAG peptide.

7.4 *in-vitro* transcription using scaffold assembly

For the assembly of scaffolds for transcription first a hybrid between the template strand and RNA was made by incubating 1 µM template DNA with 2 µM RNA in a buffer containing 25 mM Tris (pH 7.5), 5 mM $MgCl_2$ and 50 mM KCl at 45°C for 5 mins and then dropping the temperature by 2°C every 2 minutes until it reaches 21°C. This hybrid was used for the formation of ternary complex by incubating it at 30°C for 10 mins with *S. pombe* RNA polymerase II with 25 mM tris (pH 7.5), 50 mM KCl, 5 mM $MgCl_2$, 3% glycerol, 2% PVA, 0.5 mM DTT and 0.5 mg/ml BSA. 5 µM non-template strand bound to biotin was then added and incubated for 10mins at 37°C maintaining the same buffer as before. Biotin beads were then washed with 20 mM tris (pH 7.9), 3 mM HEPES (pH 7.9), 0.2% PVA, 3% glycerol, 60mM KCl, 0.5 mM DTT and 0.5 mg/ml BSA to remove the unbound DNA /RNA or polymerase.

Radioactive labelled 23ntd long transcript was synthesized off this scaffold by doing *in-vitro* transcription in the presence of 0.6 µM ATP and 0.13 µ for M 10 µCi [α-32P] UTP at 30°C for first 10 mins and then incubating for 5 minutes further after adding 5 µM ATP and UTP. Transcription were carried out in a buffer containing 20 mM tris (pH 7.9), 3

mM HEPES (pH 7.9), 8 mM $MgCl_2$, 3% Glycerol, 2% PVA, 60mM KCl, 0.5mM DTT and 0.5 mg/ml BSA. This 23-nucleotide long, elongating transcript was then chased to get longer transcripts by carrying out in-vitro transcription in the presence of all 4 nucleotides at a 7 µM concentration. Effect of addition of elongation factors like Ell1 was studies by adding it the assay. Reactions were stopped after incubation at 30 °C for the times indicated in the figures, and transcription products were purified and resolved on 15% polyacrylamide gels containing 7 M urea, 45 mM Tris-borate, and 1 mM EDTA (pH 8.3) and detected with a Molecular Dynamics Typhoon PhosphoImager.

DNA sequences (ordered from IDT):

Template strand: CTACGGTTAAGCTCACGGTACATTTCTGAATTAAGCATCATGG

RNA: ACUCUCAUGUCUGAUGCUUA

Non-template strand: ATCAGAAATGTACCGTGAGCTTAACCGTAG (With 3' Biotin tag)

7.5 Chromatin immunoprecipitation (ChIP)

S. pombe strains were grown in 200 ml YES to an OD_{600} between 0.8 and 1.0. Cells were crosslinked by adding formaldehyde to a final concentration of 1% for 15 minutes at room temperature. The crosslinking reaction was quenched with the addition of 10 ml of 2.5 M glycine for 15 minutes. Cells were harvested by centrifugation and the cell pellets were washed twice with ice-cold Phosphate-buffered saline (PBS). Cells were resuspended in 2 ml FA lysis SDS buffer (40 mM HEPES/KOH (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 % Triton X-100, 0.1% sodium deoxycholate, 0.2% SDS and protease

inhibitor). Equal volume of acid washed glass beads (Sigma G8772) were added and cells were disrupted by vortexing using a Vortex-Genie with a Turbomix attachment for 60 minutes at 4°C. Lysates were recovered from the beads and combined in a 15-ml centrifuge tube and the total volume increased to approximately 2 ml with FA-lysis SDS buffer. Lysates were sonicated at 4°C using a probe sonicator (11 cycles 10 seconds on, 30 seconds off) to generate chromatin fragments in the range 200-800 bp. Cellular debris was removed by centrifuging the sonicated extracts in 1.5 ml microfuge tubes at 4°C for 20 min at 14000 rpm.

Input DNA:

An aliquot of chromatin extract was used to prepare DNA to check the degree of chromatin shearing. 60 µl of chromatin extract was combined with 140 µl TE, pH 8.0 (10 mM Tris-HCl, 1 mM EDTA), 200 µl proteinase K buffer (50 mM Tris-HCl, 12.5 mM EDTA, 300 mM NaCl, 1% SDS) and 2 µl proteinase K (20 mg/ml New England Biolabs), and incubated at 55°C for 1 hour and then at 65°C overnight to reverse the crosslinks. Contaminating RNA was removed by treatment with 2 µl RNaseA (5 mg/ml), followed by DNA purification by phenol/chloroform extraction and ethanol precipitation using 2 µl GlycoBlue (Ambion) as a carrier. DNA was resuspended in TE to a final volume of 30 µl, 10 µl of which was run out on a 1% agarose gel. Majority of the DNA fragments were between 200 bp and 800 bp.

Chromatin Immunoprecipitations:

Usually 10 µg of antibody to be used was incubated with appropriate dynabeads (50 µl) for 2 hours at 4°C and then washed with FA-lysis buffer (40 mM HEPES/KOH (pH 7.5), 150 mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate). 400 µl

chromatin solution was mixed with 400 µl FA-lysis buffer containing AeBSF and 80 µl 10% sarkosyl. Extracts were incubated with Dynabeads (prebound with the antibodies) overnight at 4°C. Beads were concentrated with a magnetic particle concentrator, the supernatant removed and the beads washed with:

- FA lysis buffer (150 mM NaCl) for 10 minutes
- FA lysis buffer (500 mM NaCl) for 10 minutes (twice)
- TEL buffer (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.0) for 10 minutes
- TE pH 8.0 for 5 minutes twice

Bound complexes were eluted from the beads by incubating with 400 µl of elution buffer (10 mM TRIS-HCl pH 8.0, 1 mM EDTA, 250 mM NaCl, 1% SDS) at 65°C for 30 minutes. 2 µl Proteinase K (10 mg/ml) was added for digesting the protein and left overnight for reverse crosslinking at 65°C. After reversing the crosslinks, DNA was purified by phenol/chloroform extraction followed by ethanol precipitation with 2 µl GlycoBlue (Ambion) as a carrier. DNA pellets were resuspended in 60 µl TE and analyzed by either sequencing or real time PCR.

7.5.A ChIP seq

The libraries for the input and the immunoprecipitated DNA were synthesized in the Molecular Biology Core of Stowers Institute and size selected using Pippin Prep®. The libraries were then pooled and run on HiSeq machine with single read lengths of 50bp. The reads were aligned to the reference genome

"Schizosaccharomyces_pombe.ASM294v2.27. gff3" using bowtie2. Samtools was used to sort and index the Bam files. Low quality and true multi-mapped reads were filtered out with a MAPQ > 10 using samtools view. Mac2 was used to call peaks using default parameters (50>FC>5; q value<0.01), and bedtools intersect was used to assign peaks to the nearest gene.

7.5.B Real time PCR

Quantitative real time polymerase chain reactions (qPCR) were done using iQ SYBR Green Supermix (Bio-Rad) in 96 wells plates sealed with Microseal "B" sealing film (Bio-Rad). 5µl of diluted template (input or CHIP sample) was used in each 25µl reaction, and reactions were done in duplicate. Reactions were cycled on MyiQ thermocyclers (Bio-Rad), and results were analyzed by either using the percent input method or as fold enrichment over nonspecific region (ΔΔCt method).

7.6 Gene expression analysis

7.6.A RNA isolation

RNA was prepared by hot phenol extraction method. 200 ml cultures were grown to mid log phase (OD600 = 0.8-1.0), harvested on ice and were washed in DEPC water. Cells were then pulverized in liquid nitrogen using a mortar and pestle and re-suspended in 10 ml 50mM sodium Acetate (NaOAc (pH5.2)) with 1% SDS. After addition of 10 ml phenol: chloroform 5:1 (equilibrated with NaOAc (pH 5.2)), samples were vortexed for approximately 5 seconds in a fume hood and incubated at 65°C in a water bath for 2

minutes. Samples were incubated on ice for 1 minutes, vortexed for 20 seconds and centrifuged at 7500g for 5 minutes. The aqueous phase was transferred to a new tube containing 10 ml phenol: chloroform. The phenol extraction step was repeated 4 more time and then the aqueous phase (now ~4-6 ml) was mixed with 2.5 volumes 100 % ethanol and 1/10th volume 3M NaOAc and RNA precipitated using standard procedures. 100 µg of the RNA was purified further using an RNeasy mini spin column (Qiagen) per the manufacturer's instructions. The quality of the RNA was then checked on bioanalyzer.

7.6.B Library preparation

The Molecular Biology Core of the Stowers Institute for Medical Research prepared the mRNA libraries following the Illumina TruSeq Stranded mRNA sample preparation protocol using the TruSeq Standard mRNA LT Sample Prep Kit. The polyA+ RNA was purified using oligo dT magnetic beads, while the Ribo-depleted RNA was purified using Epicentre Ribo-Zero Magnetic Gold Kit (for Yeast). The purified mRNAs were fragmented, randomly primed and reverse transcribed to generate first and second strand cDNAs. The 3'-ends of the cDNAs were adenylated, adaptors ligated to the ends and the final cDNA enriched through PCR. The library was then sized selected using the Pippen Prep and validated using Agilent 2100 Bioanalyzer. The barcoded libraries were pooled together and loaded onto a flow-cell for sequencing to generate 50 bp single reads.

7.6.C RNA seq

The reads obtained were aligned to the reference genome using tophat. Samtools was used to sort and index bam files. The raw counts were obtained and then analyzed using edgeR to obtain differentially expressed genes.

7.7 Precision Run-On sequencing

PRO-seq was performed as described in (Kwak et al., 2013)and (Booth et al., 2016).

7.7.A Nuclear Run-on and RNA extraction

Strains were grown to mid-log phase in rich media (YES). 10 ml cultures of equal cell concentration were spun down and 10% equivalent of *S. cerevisiae* cells were added as spike-in. Cells were then washed in cold H20 and then resuspended in 10 mL 0.5% sarkosyl at 4°C, and incubated on ice for 20 min. Cells were then spun at a reduced RCF (400g) for 5 min at 4°C. Yeast pellets were resuspended in 120 µL of 2.5× transcription buffer (50 mM Tris –HCl, pH 7.7, 500 mM KCL, 12.5 mM MgCl2) with 6 µL 0.1 M DTT and 3.75 µL of each 1 mM biotin-NTP and the volume was brought to 285 µL with DEPC-treated H20. Finally, 15 µL 10% sarkosyl was added, and the reaction was placed at 30°C and allowed to run on for 5 min. RNA was extracted using a hot phenol approach after the run-on reaction, cells were pelleted at 400g for 5 min at 4°C and quickly resuspended in 500 µL acid phenol. An equal volume of AES buffer (50 mM NaAc pH 5.3, 10 mM EDTA, 1% SDS) was added and placed for 5 min at 65°C with periodic vortexing, followed by 5

min on ice. Two hundred microliters chloroform was added and mixed followed by centrifugation at 14,000g for 5 min (4°C). To the aqueous layer, 3 M NaOAc was added followed by ethanol precipitation with a 3× volume of 100% ethanol. The RNA pellet was air dried before being resuspended in 20 µL DEPC-treated water.

7.7.B Library preparation

The RNA was heat denatured at 65°C for 40 secs. Denatured RNA samples were then fragmented by addition of 5 µl of 1N NaOH and incubated on ice for 10 min, and neutralized by adding an equal volume of 1 M Tris-HCl pH 6.8. Buffer exchange was performed using P-30 columns following manufacturer's protocol to remove excess salt and residual NTPs. The sample was collected in 50 µL, 10mM Tris buffer and 1 µL of RNase inhibitor is added.

Prior to enriching for biotinylated RNAs, 90 µl of streptavidin-coated magnetic beads per library were washed once in buffer containing 0.1 N NaOH and 50 mM NaCl, and twice in 100 mM NaCl. Beads were then resuspended in 150 µl binding buffer (10mM Tris-HCl pH7.4, 300mM NaCl, 0.1% Triton X-100) and divided into 3 aliquots. Biotinylated RNA transcripts were then enriched by binding 50 µL of fragmented RNAs to 50 µL of pre-washed streptavidin M280 beads. Sample was incubated for 20 min at room temperature on a rotator. Bead-bound RNA samples were then washed twice in ice-cold high salt wash buffer (50 mM Tris-HCl pH 7.4, 2M NaCl, 0.5% Triton X-100), twice in binding buffer, and once in low salt buffer (5 mM Tris-HCl pH 7.4, 0.1% Triton X-100). Finally, RNAs were extracted from the beads using Trizol and ethanol precipitation.

The 3’ RNA adaptor (5Phos/rGrArUrCrGrUrCrGrGrArCrUrGrUrArGrArArCrUrCrUrGrArArC/3’Inverted dT) was ligated to the fragmented RNAs by resuspending RNA in 4 µl 3’ RNA adaptor dilution (0.5 µl of 100 µM 3’ RNA adaptor in 3.5 µl of DEPC water) and heat denaturing at 65 °C for 20 s. Denatured RNA samples were mixed with 6 µl RNA ligation mix (1 X T4 RNA ligase buffer, 1 mM ATP, 10 % PEG, 4 u/µl RNase inhibitor and 1 u/µl T4 RNA ligase I) and incubated at 20 °C for 4 h. To remove excess adaptors and salts from the ligated RNA, second biotin enrichment and purification were performed using streptavidin beads as described previously.

RNA pellets were resuspended in 5 µl DEPC water, heat denatured briefly at 65 °C for 20 s. 5’ end of RNA was then modified by mixing RNA with 5 µl 5’cap repair enzyme mix (1X ThermoPol Reaction Buffer, 2 u/µl RNase inhibitor, 0.5 u/µl RNA 5’ pyrophosphohydrolase, RppH) and incubating at 37°C for 1 h. Decapped RNAs were then mixed with 90 µl PNK mix (1 X PNK buffer, 1 mM ATP, 1 u/µl RNase inhibitor and 0.25 u/µl T4 Polynucleotide Kinase) and incubated at 37°C for 1h. The 5’end modified RNAs were then purified using Trizol LS followed by ethanol precipitation. To ligate the 5’ adaptor (rCrCrUrUrGrGrCrArCrCrCrGrArGrArArUrUrCrCrA), RNAs were dissolved in 4 µl of 5’ RNA adaptor dilution (0.5 µl of 100 µM 3’ RNA adaptor in 3.5 µl of DEPC water) and heat-denatured at 65°C. RNAs were then mixed with 6 µl RNA ligation mix (1 X T4 RNA ligase buffer, 1 mM ATP, 10 % PEG, 4 u/µl RNase inhibitor and 1 u/µl T4 RNA ligase I) and incubated at 20°C for 4 h. The cloned RNA products were then subjected to another round of biotin enrichment and purification as before and resuspended in 10 µl of DEPC water. To reverse transcribe the cloned products, RNAs were mixed with 2.5 µl RT primer mix (2.5 µM RP1 primer (AAT GAT ACG GCG ACC ACC GAG ATC TAC ACG TTC AGA GTT

CTA CAG TCC GA), 625 µM dNTP mix) and incubated at 70°C for 2 min and chilled on ice for 2 min. 6 µl of RT buffer mix (1 X First strand buffer, 5 mM DTT, 2 u/µl RNase inhibitor) was added to each RNA-primer mix and incubated at 37°C for 5 min. 1.5 µl Superscript III reverse transcriptase was mixed with the RNAs and incubated for 15 min at 45 °C, 40 min at 50 °C, 10 min at 55 °C and finally 15 min at 70 °C. 6 µl of DEPC water was then added to the RT reaction.

An aliquot of the 26 µl cDNA from each replicate was used to carry out trial amplifications to determine the optimal number of cycles to avoid over amplification of the library. Primers RPI4 (RNA PCR Primer Index 4, CAA GCA GAA GAC GGC ATA CGA GAT TGG TCA GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA), RPI5 (RNA PCR Primer Index 5, CAA GCA GAA GAC GGC ATA CGA GAT CAC TGT GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA), RPI6 (RNA PCR Primer Index 6 CAA GCA GAA GAC GGC ATA CGA GAT ATT GGC GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA) and RPI7 (RNA PCR Primer Index 7, CAA GCA GAA GAC GGC ATA CGA GAT GAT CTG GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA) were used to barcode the cDNA samples. Final full-scale PCR amplifications were carried out by mixing the remaining 24 µl cDNA samples with 0.5 µl RPI-n (25 µM) pimer, 25.5 µl full scale primer mix (1 X HF buffer, 1 M Betaine, 250 µM each dNTP mix, 250 nM RP1 primer, 0.04 u/µl Phusion DNA polymerase), and running samples using the following thermal conditions: 95 °C for 2 min, 5 cycles of 95 °C, 56°C, 72 °C each for 30 s, 13 cycles of 95 °C, 65 °C, 72 °C each for 30 s, followed by 72°C for 10 min. PCR products were purified using ethanol precipitation and resuspended in 32 µl of water. PCR amplicons running from 140 bp to 350 bp were selected using the 2% Agarose Gel Cassette (Cassette type: 2% DF Marker L) on the Pippen Prep™ (Sage Science, Software: v.5.8) instrument. Samples were then quantified using the bioanalyser.

Equimolar concentrations of library fractions were pooled together and sequenced using a mid-output flow cell on the Illlumina NextSeq platform.

7.7.C PRO-Seq analysis

The raw sequences obtained were first processed to remove the adaptor sequences and then trimmed further to a maximum length of 36 nucleotides (minimum: 15 nt). The reads then aligned to the pombe rRNA and cerevisiae rRNA were removed and rest of the reads were aligned using Bowtie to a combined genome consisting of all chromosomes from both *S. pombe* and *S. cerevisiae* facilitating the removal of reads with an ambiguous origin. Samtools was used to sort and index joint bam files and the joint reads were split by species. Bedgraph files were created by recording only the most 3' base of each read, which represents the position of the Pol II active site. The counts at each position in the bedgraph file were normalized based on the relative amount of reads aligning to the spike-in genome.

Appendix A: *S. pombe* Strains used in this study

972	*h-*
Med7-Flag (972)	*med7+::3Xflag natMX6 h-*
Med4- Flag (972)	*med4+::3Xflag natMX6 h-*
Med15- Flag (972)	*med15+::3Xflag natMX6 h-*
Med2- Flag (972)	*med2+::3Xflag natMX6 h-*
Med7 TAP med13ΔMed27Δ	*med7+::TAP-kanMX6 Δmed13::kanMX4 Δmed27::natMX6*
Med7 TAP med13ΔMed2Δ	*med7+::TAP-kanMX6 Δmed13::kanMX4 Δmed2::natMX6*
Ell1-FLAG (972)	*ell1+::3Xflag natMX6 h-*
Eaf1-FLAG (972)	*eaf1+:: 3Xflag natMX6 h-*
ED666	*From Bioneer Inc. h+*
Pem2	*From (Roguev et al., 2007b) h-*
ell1Δ (Pem2)	*Δell1::natMX6 h-*
eaf1Δ (Pem2)	*Δeaf1::natMX6 h-*
ebp1Δ (Pem2)	*ΔSPAC6G9.15c::natMX6 h-*
brl1Δ (Pem2)	*Δbrl1::kanMX4 h-*
brl1Δ ell1Δ (Pem2)	*Δbrl1::kanMX4 Δell1::natMX6 h-*
brl1Δ eaf1Δ (Pem2)	*Δbrl1::kanMX4 Δeaf1::natMX6 h-*

brl1Δ ebp1Δ (Pem2)	*Δbrl1::kanMX4 ΔSPAC6G9.15c::natMX6 h-*
Htb1-K119R (Ub mutant)	*otr1R(Sph1)::ura+ htb1::htb1-K119R kanMX6 (from Shiv Grewal)*

Appendix B: Primers used in this study

For strain construction:

Med4_Forward	*ATAAAGAAGTAGAATCGCCAGCTAATAAGGATGTCTTTGCAGGATTTGATCTTTTTGATCCAGAAATGGAAGAAGATTTTCGGATCCCCGGGTTAATTAA*
Med4_Reverse	*GGTATTTGGTCAAAGCAAAAAACTCTATAATTAGAAAAGTAAAATACCAGATATCACTTGACAAATAACACAATTTAGAAGAATTCGAGCTCGTTTAAAC*
MED15_Forwar	*TAAAGCAGCTTCAAGTGGGTAACGAAGATGAAGACAACATCGCATTTCCAACATCTACAAACATATGGCAAGTTGTGATTCGGATCCCCGGGTTAATTAA*
MED15_Revers	*CAAATATGTCCAATTTGCTTTTTGAGACGACGTGCAATAAATAAATAAAAAACAATTGCGAACCATCGTATCACAAAATTGAATTCGAGCTCGT*
MED2_Forward	*TAAATCTTCCTACAGACATTCCAACCACTGAAAGCATAAATGACTTGTTTGGTGAGAACTTTGACTTCACCATGACAAAACGGATCCCCGGGTT*
Med2_del_F	*CATCCTATTTAACTAAAAGGTTGTTCAGATCTATTCAGCTTTGTGATTGCATAGAACTCCCATTGTTTGTTATTCCTTCC*
Med2_Reverse	*TAATACATGAAAAAAACATTGACAAATTGAAACAACCACGTTAAACCGATCATACATAGCTGAACTGAAGTCAACTCAACGAATTCGAGCTCGT*
Med27_del_F	*CCATCAATATTTATTTAATATACGCATTGTTAATTTTTTTTTCATAATTATTTTAAGTTTGTTAAGGCTACCAGAAATGCGGATCCCCGGGTTAAT*
Med27_Revers	*AATTGTAACTCAAGCTCAAAACGAAGCAAAAATGGCAATTATGAACGTCTTATTTCATTTGAATATCAAAATTAGTGTTGAATTCGAGCTCGTTT*
Ell1_Forward	*AAGATGAACTTATATCTCTACATTCTCAGCTCAAAAGCTGGAAAAATACACTTTACGATGCTTCCTCGGAGCTAGCCCTCCGGATCCCCGGGTTA*
Ell1_Del_F	*CTTAAACAAGCTTCTCCGTTGTGCCATCTAGCTAATATAATCATTTTGAGAGGCTTTTACTATCGATCTATTTGGGTTGACGGATCCCCGGGTTA*
Ell1_Reverse	*AAGCTGAGATAACCGTCACCTTTATAAAACATTGGAAATAAAACTGAGTGTTAAGAAAACCGGAAATTTATTGAGAAAAGGAATTCGAGCTCG*
Eaf1_Forward	*GGGGTCTATCTTCGCAAGAGAGGGATTATGCTTCTTCTGCTCAGGCAGAGGGTATCAGCAGCGCTTCCGAGGATGAGGATCGGATCCCCGGG*
Eaf1_Del_F	*TCTTACCTTACGTTATTTATTTGATTTATATCGAAATTTCCAATTTCGTACAGGCCTGACTTTTACCATTATAAACAATCCGGATCCCCGGGTTAAT*
Eaf1_Reverse	*TTTCGTTGCAAGTCGTTCATTAACAGATTTTCTCATTATCGTGTTTGGAAAGTGTATACGAAACACAGCATAAAGCATACGAATTCGAGCTCGTT*
Ebp1_Del_F	*CGAAATCTTTTAATTCCATAGCATATCTCTTTTTTAAGTAAATACTTTTATTTGAATAGTTTATACAGCGCATGCGATGACGGATCCCCGGGTTAA*
Ebp1_reverse	*AACTTTATTAGTGAAAGAATGACATTCTGTTATTGAAGACCGACTCAAATCTTTCAAGTTATGGGAGTTAATACAATTATGAATTCGAGCTCGTT*

For qPCR:

Tdh promoter_F	*GTCGGTGTCAACGAGGAGAA*
Tdh promoter_R	*AGTGGTCATGAGACCCTCCT*
Tdh 3'_F	*TCGTCAAGCTCGTCTCTTGG*
Tdh 3'_R	*GCAGTGTAGGCAACCAAGTC*
Chr1_left_F	*CCTGCTTACCCTTGGACTTT*
Chr1_left_R	*CTTCACCACCATTTCGCTTTAC*
Chr1_left _F	*CTCCCAAACTCCACACCATATC*
Chr1_left _R	*CAGTTGCGAGTCCTGACAATAG*
Chr1_right_F	*AGCCTGAAACCAGCAATCA*
Chr1_right_R	*TTAGGTCCGGCCAACAATAC*
Chr1_right _F	*TGGCCGATTGGCTTAAAGTAG*
Chr1_right _R	*GTAGGCGAAGAACCACCTAATG*
Chr2_left_F	*GTCCTGGGATTAGTGGTTTGAT*
Chr2_left_R	*CTTCACCACCATTTCGCTTTAC*
Chr2_left _F	*ACTGAGTTTGGGTAGACTTGTATT*
Chr2_left _R	*CAACCCTTCTGACATTCGTTTATT*
Chr2_right_F	*GTCCTGGGATTAGTGGTTTGAT*
Chr2_right_R	*CTTCACCACCATTTCGCTTTAC*
Chr2_right_F	*CCTAGAGCAGGGACTTGATTTG*
Chr2_right_R	*TGGTGTTGGTAAAGAGGGATTG*
Chr1_control_F	*GCTCCTTATCCCAATCCTTACC*
Chr1_control_R	*CCTATCCTCACCACCATCAAC*
Chr2_control_F	*GGTTCCTCCGTGTTCGATATT*

Appendix C: SGA results

id	gene	product	ell1Δ			eaf1Δ			ebp1Δ			ell1Δ_mpa		
SPAC22A12.07c	ogm1	protein O-mannosyltransferase Ogm1	1	1	0	1	1	1	1	1	1	1	1	0
SPCC1919.15	brl1	ubiquitin-protein ligase E3 Brl1	1	1	1	0	1	1	0	1	0	1	1	1
SPAC31G5.03	rps1101	40S ribosomal protein S11 (predicted)	1	1	1	0	0	0	1	0	1	1	1	1
SPAC6F6.12	atg24	autophagy associated protein Atg24 (predicted)	1	1	0	1	1	0	0	1	1	1	1	0
SPCC794.08	efr3	HEAT repeat protein involved in protein localization to plasma membrane	1	1	1	0	1	1	0	1	1	0	1	0
SPBC1105.10	rav1	RAVE complex subunit Rav1	1	1	0	1	1	1	0	1	0	1	1	0
SPAC3H5.13	new4	conserved eukaryotic protein	1	1	1	0	0	1	0	1	0	1	1	1
SPAC3G6.02	rpn15	proteasome regulatory particle, lid subcomplex subunit Rpn15/Dss1	1	1	1	0	1	0	0	0	1	1	1	1
SPAC1486.04c	alm1	medial ring protein Alm1	1	0	0	0	0	1	1	1	1	1	1	0
SPAC6G10.12c	ace2	transcription factor Ace2	1	1	0	0	0	0	1	1	1	1	1	0
SPBC4F6.06	kin1	microtubule affinity-regulating kinase Kin1	1	1	1	0	1	1	0	1	1	0	0	0
SPBC577.15c	sim3	NASP family CENP-A chaperone	1	1	1	0	1	0	0	1	0	1	1	0
SPCC306.09c	cap1	adenylyl cyclase-associated protein Cap1	1	1	1	0	1	1	0	0	1	1	0	0
SPBC3D6.08c	lsm1	mRNA decapping complex subunit (predicted)	1	1	1	0	1	0	0	0	1	1	1	0
SPAC8C9.07	fyv7	rRNA processing protein Fyv7 (predicted)	1	1	1	0	0	0	0	0	1	1	1	1
SPCC285.13c	nup60	nucleoporin Nup60	1	1	0	1	1	1	0	0	0	1	1	0
SPBC1A4.09	pus7	pseudouridine synthase Pus7 (predicted)	1	1	1	1	0	0	0	0	0	1	1	1
SPBC215.08c	arg4	arginine specific carbamoyl-phosphate synthase Arg4	1	1	1	0	1	1	0	0	0	1	0	1
SPCC830.06	cnb1	calcineurin regulatory subunit (calcineurin B)	1	1	1	0	1	0	0	0	0	1	1	1
SPBC21H7.04	dbp7	ATP-dependent RNA helicase Dbp7 (predicted)	1	1	1	0	0	1	0	0	0	1	1	1
SPAPB1E7.11c		Schizosaccharomyces specific protein	1	1	0	0	0	0	1	1	1	1	0	0
SPAC3H5.10	rpl3202	60S ribosomal protein L32 (predicted)	1	1	0	0	0	0	1	1	0	1	1	0
SPBC16C6.01c		lysine methyltransferase (predicted)	0	0	1	0	0	0	1	0	1	1	1	1
SPAC3G9.08	png1	ING family homolog Png1	1	1	0	0	0	0	0	1	1	1	1	0
SPAC15E1.05c	ect1	ethanolamine-phosphate cytidylyltransferase (predicted)	1	1	0	0	0	0	0	1	1	1	1	0
SPBC530.06c	clu1	clustered mitochondria (cluA/CLU1) homolog Clu1 (predicted)	1	1	1	0	1	1	0	1	0	0	0	0
SPAC26A3.04	rpl2002	60S ribosomal protein L20 (predicted)	1	1	0	0	1	0	0	1	0	1	1	0
SPAP7G5.03	prm1	conjugation protein Prm1	1	1	0	0	0	1	0	1	0	1	1	0
SPAP7G5.05	rpl1002	60S ribosomal protein L10	1	1	0	0	0	1	0	1	0	1	1	0
SPBC13G1.14c		RNA-binding protein (predicted)	1	1	1	0	0	0	0	1	0	1	1	0
SPBC577.05c	rec27	meiotic recombination protein Rec27	1	1	1	1	0	0	0	0	1	1	0	0
SPAC607.02c	SPAC607.02c	conserved fungal protein	1	0	0	0	1	1	0	0	1	1	1	0
SPAC926.03	rlc1	myosin II regulatory light chain Rlc1	1	1	0	0	0	1	0	0	1	1	1	0
SPBC21C3.12c		DUF953 thioredoxin family protein	1	1	0	0	0	0	0	0	1	1	1	1
SPCC1020.01c	pma2	P-type proton ATPase, P3-type Pma2	0	1	1	1	1	0	0	0	0	0	1	1
SPBC30D10.10c	tor1	phosphatidylinositol kinase Tor1	1	0	1	1	0	0	0	0	0	1	1	1
SPBC2G2.03c	sbh1	translocon beta subunit Sbh1 (predicted)	1	1	0	0	1	1	0	0	0	1	1	0
SPBC216.03		conserved fungal protein	1	1	0	0	1	0	0	0	0	1	1	1
SPBC1A4.03c	top2	DNA topoisomerase II	0	1	1	0	1	0	0	0	0	1	1	1
SPAC31F12.01	zds1	zds family protein phosphatase type A regulator Zds1 (predicted)	1	1	1	0	0	1	0	0	0	1	1	0
SPBC19G7.04		HMG box protein	1	1	1	0	0	1	0	0	0	1	1	0
SPCC1753.05	rsm1	RNA export factor Rsm1	1	1	1	0	0	1	0	0	0	1	1	0
SPBC2F12.11c	rep2	MBF transcription factor activator Rep2	1	1	0	0	0	1	0	0	0	1	1	1
SPAC3H1.09c	avt3	vacuolar amino acid transmembrane transporter Avt3	0	1	1	0	0	1	0	0	0	1	1	1
SPBC21C3.13	rps1901	40S ribosomal protein S19 (predicted)	1	1	1	0	0	0	0	0	0	1	1	1
SPAC57A7.04c	pabp	mRNA export shuttling protein	1	1	1	0	0	0	0	0	0	1	1	1
SPCC736.11	ago1	argonaute	1	1	1	0	0	0	0	0	0	1	1	1

SPCC1450.03		*ribonucleoprotein (RNP) complex (predicted)*	1	1	1	0	0	0	0	0	0	1	1	1
SPBC16E9.18	***psd1***	*phosphatidylserine decarboxylase Psd1*	1	1	1	0	0	0	0	0	0	1	1	1
SPCC1739.06c		*uroporphyrin methyltransferase (predicted)*	1	1	1	0	0	0	0	0	0	1	1	1
SPAC22F3.12c	***rgs1***	*regulator of G-protein signaling Rgs1*	1	1	1	0	0	0	0	0	0	1	1	1
SPCC1682.13	***laf2***	*Clr6 associated factor 2, Laf2*	1	1	1	0	0	0	0	0	0	1	1	1
SPBP23A10.14c	***ell1***	*RNA polymerase II transcription elongation factor SpELL*	1	1	1	0	0	0	0	0	0	1	1	1
SPBC16E9.07	***mug100***	*Schizosaccharomyces pombe specific protein Mug100*	1	1	1	0	0	0	0	0	0	1	1	1
SPCC584.03c		*Ran GTP-binding protein (predicted)*	1	1	1	0	0	0	0	0	0	1	1	1
SPCC576.12c	***mhf2***	*CENP-X homolog, FANCM-MHF complex subunit Mhf2*	1	1	1	0	0	0	0	0	0	1	1	1
SPCC962.01		*C2 domain endoplasmic reticulum membrane organization protein (predicted)*	1	1	1	0	0	0	0	0	0	1	1	1
SPBC839.15c	***tef103***	*translation elongation factor EF-1 alpha Ef1a-c*	1	1	1	0	0	0	0	0	0	1	1	1
SPBC713.08	***mim1***	*mitochondrial TOM complex assembly protein Mim1 (predicted)*	1	1	1	0	0	0	0	0	0	1	1	1
SPCC16C4.03	***pin1***	*peptidyl-prolyl cis-trans isomerase Pin1*	1	1	1	0	0	0	0	0	0	1	1	1
SPBC11G11.05	***rpa34***	*DNA-directed RNA polymerase I complex subunit Rpa34 (predicted)*	1	1	1	0	0	0	0	0	0	1	1	1
SPAC6G9.14		*RNA-binding protein (predicted)*	0	1	0	0	0	0	1	1	1	0	1	0
SPAC6G9.15c	***Ebp1***	*Schizosaccharomyces specific protein*	0	0	1	0	0	0	1	1	1	0	1	0
SPCC1739.12	***ppe1***	*serine/threonine protein phosphatase Ppe1*	0	0	1	0	1	0	1	1	0	0	0	1
SPAC11E3.01c	***swr1***	*SNF2 family ATP-dependent DNA helicase Swr1*	1	1	0	0	0	0	1	1	0	1	0	0
SPAC16E8.01	***shd1***	*cytoskeletal protein binding protein Sla1 family, Shd1 (predicted)*	1	1	0	0	0	0	1	0	0	1	1	0
SPCC1739.01		*zf-CCCH type zinc finger protein*	1	1	0	0	0	0	0	1	1	1	0	0
SPBC1D7.04	***mlo3***	*RNA binding protein Mlo3*	1	1	1	0	0	1	0	1	0	0	0	0
SPCC24B10.12	***cgi121***	*EKC/KEOPS complex subunit Cgi121 (predicted)*	1	1	0	0	0	0	0	1	0	1	1	0
SPBC2D10.18	***abc1***	*ABC1 kinase family ubiquinone biosynthesis protein Abc1/Coq8*	1	1	1	0	0	1	0	0	1	0	0	0
SPAC227.07c	***pab1***	*protein phosphatase regulatory subunit Pab1*	1	1	0	0	0	1	0	0	1	0	1	0
SPAC15A10.11	***ubr11***	*UBR ubiquitin-protein ligase E3 Ubr11*	1	0	1	0	0	1	0	0	1	1	0	0
SPAC15A10.03c	***rad54***	*DNA-dependent ATPase Rad54/Rhp54*	1	1	1	0	0	0	0	0	1	1	0	0
SPAC15A10.16	***bud6***	*actin interacting protein 3 homolog Bud6*	1	1	0	0	0	0	0	0	1	1	1	0
SPAC1071.02	***mms19***	*Dos2 silencing complex subunit Mms19*	1	1	0	0	0	0	0	0	1	1	1	0
SPBC32F12.11	***tdh1***	*glyceraldehyde-3-phosphate dehydrogenase Tdh1*	1	1	0	0	0	0	0	0	1	1	1	0
SPBC660.09	***mug168***	*Schizosaccharomyces specific protein Mug168*	1	1	0	0	0	0	0	0	1	1	1	0
SPAC22F3.10c	***gcs1***	*glutamate-cysteine ligase Gcs1*	1	1	0	0	0	0	0	0	1	1	1	0
SPAC17G6.08	***pep7***	*prevacuole/endosomal FYVE tethering component Pep7 (predicted)*	1	1	0	0	0	0	0	0	1	1	1	0
SPCC285.14	***trs130***	*TRAPP complex subunit Trs130 (predicted)*	0	1	0	1	1	1	0	0	0	0	1	0
SPCC1223.05c	***rpl3702***	*60S ribosomal protein L37 (predicted)*	0	0	1	1	1	1	0	0	0	0	0	1
SPCC1235.11	***mpc1***	*mitochondrial pyruvate transmembrane transporter subunit Mpc1 (predicted)*	1	1	0	1	1	0	0	0	0	0	1	0
SPBC1718.07c	***zfs1***	*CCCH tandem zinc finger protein, human Tristetraprolin homolog Zfs1, involved in mRNA catabolism*	1	1	0	1	0	0	0	0	0	1	1	0
SPCC74.02c	***ppn1***	*mRNA cleavage and polyadenylation specificity factor complex associated protein*	1	1	0	1	0	0	0	0	0	1	1	0
SPCC285.17	***spp27***	*RNA polymerase I upstream activation factor complex subunit Spp27*	1	1	0	1	0	0	0	0	0	1	1	0
SPBC8D2.03c	***hhf2***	*histone H4 h4.2*	1	1	0	1	0	0	0	0	0	1	1	0
SPCC364.03	***rpl1702***	*60S ribosomal protein L17 (predicted)*	1	1	0	1	0	0	0	0	0	0	1	1
SPAC30C2.06c	***dml1***	*mitochondrial inheritance GTPase, tubulin-like (predicted)*	1	0	0	0	1	1	0	0	0	1	1	0
SPCC5E4.07	***rpl2802***	*60S ribosomal protein L27/L28*	1	0	0	0	1	1	0	0	0	1	1	0

SPCC11E10.04	***ppr6***	*mitochondrial PPR repreat protein Ppr6*	0	1	0	0	1	1	0	0	0	0	1	1
SPBC14C8.17c	***spt8***	*SAGA complex subunit Spt8*	1	1	0	0	1	0	0	0	0	1	1	0
SPCC1322.10		*cell wall protein Pwp1*	1	1	0	0	1	0	0	0	0	1	1	0
SPBC19C2.14	***smd3***	*Sm snRNP core protein Smd3*	1	1	0	0	1	0	0	0	0	1	1	0
SPAC1556.01c	***rad50***	*DNA repair protein Rad50*	1	1	0	0	1	0	0	0	0	1	1	0
SPAC458.05	***pik3***	*phosphatidylinositol 3-kinase Pik3*	1	1	0	0	1	0	0	0	0	1	1	0
SPBC365.16		*conserved protein*	1	0	1	0	1	0	0	0	0	1	1	0
SPCC1259.01c	***rps1802***	*40S ribosomal protein S18 (predicted)*	1	0	1	0	1	0	0	0	0	0	1	1
SPBC646.12c	***gap1***	*GTPase activating protein Gap1*	0	1	1	0	1	0	0	0	0	0	1	1
SPAC1F7.01c	***spt6***	*nucleosome remodeling protein Spt6*	0	1	1	0	1	0	0	0	0	0	1	1
SPAC1556.08c	***cbs2***	*AMP-activated protein kinase gamma subunit cbs2*	1	1	0	0	0	1	0	0	0	1	1	0
SPAC6B12.15	***cpc2***	*RACK1 ortholog Cpc2*	1	1	0	0	0	1	0	0	0	1	1	0
SPBC651.09c	***prf1***	*RNA polymerase II associated Paf1 complex (predicted)*	1	1	0	0	0	1	0	0	0	1	1	0
SPBC4B4.07c	***usp102***	*U1 snRNP-associated protein Usp102*	1	1	0	0	0	1	0	0	0	1	1	0
SPAC3A12.10	***rpl2001***	*60S ribosomal protein L20a (predicted)*	1	1	0	0	0	1	0	0	0	1	1	0
SPAC57A7.13		*RNA-binding protein, involved in splicing (predicted)*	0	1	1	0	0	1	0	0	0	0	1	1
SPBC32F12.05c	***cwf12***	*complexed with Cdc5 protein Cwf12*	0	1	1	0	0	1	0	0	0	0	1	1
SPBC56F2.01	***pof12***	*F-box protein Pof12*	1	1	1	0	0	0	0	0	0	1	1	0
SPCC1795.02c	***vcx1***	*CaCA proton/calcium exchanger (predicted)*	1	1	1	0	0	0	0	0	0	1	1	0
SPCC794.10	***ugp1***	*UTP-glucose-1-phosphate uridylyltransferase-like Ugp1*	1	1	1	0	0	0	0	0	0	1	1	0
SPBC83.03c	***tas3***	*RITS complex subunit 3*	1	1	1	0	0	0	0	0	0	1	1	0
SPAC13G7.12c	***eki1***	*choline/ethanolamine kinase Eki1 (predicted)*	1	1	1	0	0	0	0	0	0	1	1	0
SPCC11E10.08	***rik1***	*silencing protein Rik1*	1	1	1	0	0	0	0	0	0	1	1	0
SPBC1A4.04		*Schizosaccharomyces specific protein*	1	1	1	0	0	0	0	0	0	1	1	0
SPAC15A10.10	***mde6***	*Muskelin homolog (predicted)*	1	1	1	0	0	0	0	0	0	1	1	0
SPAC26H5.03	***pcf2***	*CAF assembly factor (CAF-1) complex subunit B, Pcf2*	1	1	1	0	0	0	0	0	0	1	1	0
SPAC2F7.02c	***psr1***	*NLI interacting factor family phosphatase Psr1 (predicted)*	1	1	1	0	0	0	0	0	0	1	1	0
SPAC22H10.07	***scd2***	*scaffold protein Scd2*	1	1	1	0	0	0	0	0	0	1	1	0
SPCC895.05	***for3***	*formin For3*	1	1	1	0	0	0	0	0	0	1	1	0
SPAC30D11.13	***hus5***	*SUMO conjugating enzyme E2 Hus5*	1	1	1	0	0	0	0	0	0	1	1	0
SPBC725.10		*mitochondrial transport protein, tspO homolog (predicted)*	1	1	1	0	0	0	0	0	0	1	0	1
SPBC776.17	***rrp7***	*rRNA processing protein Rrp7 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	1
SPAC23C11.02c	***rps23***	*40S ribosomal protein S23 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	1
SPBC16C6.06	***vps10***	*sorting receptor for vacuolar proteins, Vps10*	1	1	0	0	0	0	0	0	0	1	1	1
SPCC1259.05c	***cox9***	*cytochrome c oxidase subunit VIIa (predicted)*	1	1	0	0	0	0	0	0	0	1	1	1
SPBP23A10.12	***frg1***	*FRG1 family protein, involved in mRNA splicing (predicted)*	1	1	0	0	0	0	0	0	0	1	1	1
SPBC16E9.09c	***erp5***	*COPII vesicle coat component Erp5/Erp6 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	1
SPCC613.03		*conserved fungal protein*	1	1	0	0	0	0	0	0	0	1	1	1
SPAC17A5.14	***exo2***	*exonuclease II Exo2*	1	1	0	0	0	0	0	0	0	1	1	1
SPBC16E9.20		*dubious*	1	1	0	0	0	0	0	0	0	1	1	1
SPBC16E9.16c	***lsd90***	*Lsd90 protein*	1	1	0	0	0	0	0	0	0	1	1	1
SPBC16E9.15		*heat shock factor binding protein (predicted)*	1	1	0	0	0	0	0	0	0	1	1	1
SPBC29A3.14c	***trt1***	*telomerase reverse transcriptase 1 protein Trt1*	1	1	0	0	0	0	0	0	0	1	1	1
SPAC1783.04c	***hst4***	*Sirtuin family histone deacetylase Hst4*	1	1	0	0	0	0	0	0	0	1	1	1
SPAC2H10.04		*dubious*	1	0	1	0	0	0	0	0	0	1	1	1
SPAC26H5.11	***mug56***	*spore wall assembly protein Mug56 (predicted)*	1	0	1	0	0	0	0	0	0	1	1	1
SPAPB2B4.02	***grx5***	*monothiol glutaredoxin Grx5*	1	0	1	0	0	0	0	0	0	1	1	1
SPBC3H7.15	***hhp1***	*serine/threonine protein kinase Hhp1*	0	1	1	0	0	0	0	0	0	1	1	1
SPBC902.03		*Nem1-Spo7 complex regulatory subunit (predicted)*	0	1	1	0	0	0	0	0	0	1	1	1
SPAC821.05	***tif38***	*translation initiation factor eIF3h (p40)*	0	1	1	0	0	0	0	0	0	1	1	1

SPBC3B8.10c	***nem1***	*Nem1-Spo7 phosphatase complex catalytic subunit Nem1 (predicted)*	0	1	1	0	0	0	0	0	0	1	1	1
SPBC3H7.05c		*mitochondrial Membrane Bound O-Acyl Transferase (MBOAT) family*	0	1	1	0	0	0	0	0	0	1	1	1
SPCC16C4.12	***naa20***	*NatB N-acetyltransferase complex catalytic subunit Naa20 (predicted)*	0	0	0	0	0	1	1	1	1	0	0	0
SPAC1142.05	***ctr5***	*copper transporter complex subunit Ctr5*	1	0	0	0	0	0	1	1	0	1	0	0
SPAC26A3.16	***dph1***	*UBA domain protein Dph1*	1	0	0	0	0	0	1	0	1	1	0	0
SPAC6G9.04	***spo7***	*sporulation protein Spo7*	0	0	0	0	0	1	0	1	1	0	1	0
SPAC6G9.05	***pcd1***	*coenzyme A diphosphatase (predicted)*	0	0	1	0	0	0	0	1	1	0	0	1
SPBC1E8.05		*conserved fungal protein*	0	1	0	0	1	1	0	1	0	0	0	0
SPCP1E11.04c	***pal1***	*membrane associated protein Pal1*	0	1	0	0	0	1	0	1	0	0	1	0
SPBC18H10.11c	***ppr2***	*mitochondrial PPR repeat protein Ppr2*	1	1	1	0	0	0	0	0	1	0	0	0
SPBC18H10.14	***rps1601***	*40S ribosomal protein S16 (predicted)*	1	1	0	0	0	0	0	0	1	0	1	0
SPAC23A1.07		*ubiquitin-protein ligase E3 (predicted)*	1	0	0	0	0	0	0	0	1	1	1	0
SPAC18G6.04c	***shm2***	*serine hydroxymethyltransferase Shm2 (predicted)*	1	0	0	0	0	0	0	0	1	1	1	0
SPAC3H5.12c	***rpl501***	*60S ribosomal protein L5 (predicted)*	1	0	0	0	0	0	0	0	1	1	0	1
SPAC5D6.05	***med18***	*mediator complex subunit Med18*	0	1	0	0	0	0	0	0	1	1	1	0
SPCC594.05c	***spf1***	*Set1C PHD Finger protein Spf1*	0	1	0	0	0	0	0	0	1	1	1	0
SPBC409.19c		*metaxin (predicted)*	0	1	0	0	0	0	0	0	1	1	1	0
SPCC1223.13	***cbf12***	*CBF1/Su(H)/LAG-1 family transcription factor Cbf12*	1	0	0	1	1	1	0	0	0	0	0	0
SPCC297.06c		*conserved fungal protein*	0	0	0	1	1	1	0	0	0	1	0	0
SPCC737.09c	***hmt1***	*vacuolar transmembrane transporter Hmt1*	0	0	0	1	1	1	0	0	0	0	1	0
SPCC297.04c	***set7***	*histone lysine methyltransferase Set7 (predicted)*	0	0	0	1	1	1	0	0	0	0	0	1
SPCC1223.15c	***spc19***	*DASH complex subunit Spc19*	0	0	0	1	1	0	0	0	0	1	1	0
SPCC14G10.04		*Schizosaccharomyces specific protein*	1	0	0	1	0	0	0	0	0	1	0	1
SPAC15E1.07c	***moa1***	*meiotic cohesin complex associated protein (Meikin) Moa1*	0	1	1	0	1	0	0	0	0	0	1	0
SPBC4F6.10	***vps901***	*guanyl-nucleotide exchange factor Vps901 (predicted)*	0	1	1	0	1	0	0	0	0	0	1	0
SPBC725.11c	***php2***	*CCAAT-binding factor complex subunit Php2*	0	1	1	0	1	0	0	0	0	0	1	0
SPCC24B10.11c	***tho7***	*THO complex subunit Tho7 (predicted)*	1	1	1	0	0	1	0	0	0	0	0	0
SPAC57A10.02	***cdr2***	*serine/threonine protein kinase Cdr2*	1	0	1	0	0	1	0	0	0	1	0	0
SPCC736.07c	***uri1***	*unconventional prefoldin chaperone involved protein complex assembly Uri1 (predicted)*	1	0	0	0	0	1	0	0	0	1	1	0
SPBC428.08c	***clr4***	*histone H3 lysine methyltransferase Clr4*	0	1	1	0	0	1	0	0	0	0	1	0
SPAC521.03		*short chain dehydrogenase (predicted)*	0	1	1	0	0	1	0	0	0	0	1	0
SPBC776.02c	***dis2***	*serine/threonine protein phosphatase PP1 subfamily, Dis2 :*	0	1	0	0	0	1	0	0	0	1	1	0
SPCC162.12	***tco89***	*TORC1 subunit Tco89*	0	1	0	0	0	1	0	0	0	1	1	0
SPCC285.11	***dsc5***	*UBX domain containing protein required for Sre1 cleavage*	0	0	1	0	0	1	0	0	0	0	1	1
SPAC22H10.11c		*TOR signaling pathway transcriptional corepressor Crf1 (predicted)*	1	1	1	0	0	0	0	0	0	1	0	0
SPBC725.14	***arg6***	*acetylglutamate synthase Arg6 (predicted)*	1	1	1	0	0	0	0	0	0	1	0	0
SPBC839.19	***new20***	*conserved eukaryotic protein*	1	1	1	0	0	0	0	0	0	1	0	0
SPBC530.07c		*TENA/THI family protein*	1	1	1	0	0	0	0	0	0	1	0	0
SPBC609.05	***pob3***	*FACT complex subunit Pob3*	1	1	1	0	0	0	0	0	0	0	1	0
SPAC2C4.05	***cor1***	*cornichon family protein (predicted)*	1	1	1	0	0	0	0	0	0	0	1	0
SPAC17H9.13c		*glutamate 5-kinase (predicted)*	1	1	1	0	0	0	0	0	0	0	1	0
SPCC4B3.10c	***ipk1***	*inositol 1,3,4,5,6-pentakisphosphate (IP5) kinase*	1	1	1	0	0	0	0	0	0	0	1	0
SPBC3H7.03c		*2-oxoglutarate dehydrogenase (lipoamide) (e1 component of oxoglutarate dehydrogenase complex) (predicted)*	1	1	1	0	0	0	0	0	0	0	1	0
SPCC16C4.11	***pef1***	*Pho85/PhoA-like cyclin-dependent kinase Pef1*	1	1	1	0	0	0	0	0	0	0	1	0
SPCC613.06	***rpl902***	*60S ribosomal protein L9*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC9B6.07	***nop52***	*nucleolar protein Nop52 family Rrp1 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0

SPBC428.06c	***rxt2***	*histone deacetylase complex subunit Rxt2*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC1322.03	***trp1322***	*TRP-like ion channel (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC1840.08c	***pdi5***	*protein disulfide isomerase (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC23E6.10c	***mri1***	*methylthioribose-1-phosphate isomerase Mri1 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC126.15c	***sec65***	*signal recognition particle subunit Sec65 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC23C11.13c	***hpt1***	*xanthine phosphoribosyltransferase (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC18B5.03	***wee1***	*M phase inhibitor protein kinase Wee1*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC25D12.02c	***dnt1***	*nucleolar protein Dnt1*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC1235.08c	***pdh1***	*Golgi to ER protein (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC1093.01	***ppr5***	*mitochondrial PPR repeat protein Ppr5*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC2F3.12c	***plp1***	*thioredoxin fold protein Plp1 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC16E9.02c		*CUE domain protein*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC17G8.13c	***mst2***	*histone acetyltransferase Mst2*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC22E12.04	***ccs1***	*superoxide dismutase copper chaperone Ccs1*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC23E6.09	***ssn6***	*transcriptional corepressor Ssn6*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC1071.07c	***rps1502***	*40S ribosomal protein S15 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC146.13c	***myo1***	*myosin type I*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC1F12.05	***any2***	*arrestin-related endocytic adaptor Any2 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC19G7.10c	***pdc2***	*topoisomerase II-associated deadenylation-dependent mRNA-decapping factor Pdc2 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC1E8.02		*ubiquitin family protein (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC794.04c		*amino acid transmembrane transporter (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPBP23A10.10	***ppk32***	*serine/threonine protein kinase Ppk32 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC27E2.07	***pvg2***	*galactose residue biosynthesis protein Pvg2*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC18H10.02	***lcf1***	*long-chain-fatty-acid-CoA ligase Lcf1*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC794.11c	***ent3***	*ENTH/VHS domain protein Ent3 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC29B12.04	***snz1***	*pyridoxine biosynthesis protein*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC1739.14	***npp106***	*nucleoporin Npp106*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC22F3.13	***tsc1***	*hamartin*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC16E9.17c	***rem1***	*meiosis-specific cyclin Rem1*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC23E6.08	***sat1***	*Golgi membrane exchange factor subunit Sat1 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC613.11c	***meu23***	*mug2/mug135/meu2 family*	1	1	0	0	0	0	0	0	0	1	1	0
SPBP23A10.17		*conserved fungal protein*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC553.04	***cyp9***	*WD repeat containing cyclophilin family peptidyl-prolyl cis-trans isomerase Cyp9 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPCC1902.01	***gaf1***	*transcription factor Gaf1*	1	1	0	0	0	0	0	0	0	1	1	0
SPBP16F5.04	***ubc7***	*ubiquitin conjugating enzyme E2 Ubc7/UbcP3*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC17H9.10c	***ddb1***	*damaged DNA binding protein Ddb1*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC1952.11c	***ure2***	*urease Ure2*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC8C9.14	***prr1***	*transcription factor Prr1*	1	1	0	0	0	0	0	0	0	1	1	0
SPAC1952.02	***tma23***	*ribosome biogenesis protein Tma23 (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC16H5.08c		*ribosome biogenesis ATPase, Arb family ABCF2-like (predicted)*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC776.05		*transmembrane transporter (predicted)*	1	1	0	0	0	0	0	0	0	1	0	1
SPBC800.03	***clr3***	*histone deacetylase (class II) Clr3*	1	0	1	0	0	0	0	0	0	1	1	0
SPAC31G5.21		*human FAM32A homolog*	1	0	1	0	0	0	0	0	0	1	1	0
SPAC1399.03	***fur4***	*uracil permease*	1	0	1	0	0	0	0	0	0	1	1	0
SPAC22G7.05	***kri1***	*ribosome biogenesis protein Kri1 (predicted)*	1	0	1	0	0	0	0	0	0	1	0	1
SPAC1296.06	***tah18***	*NADPH-dependent diflavin oxidoreductase, involved in iron-sulfur cluster assembly Tah18 (predicted)*	1	0	1	0	0	0	0	0	0	1	0	1
SPBC16A3.18	***cip1***	*RNA-binding protein Cip1*	1	0	1	0	0	0	0	0	0	1	0	1
SPBC725.15	***ura5***	*orotate phosphoribosyltransferase Ura5*	1	0	1	0	0	0	0	0	0	1	0	1

SPCC962.04	***rps1201***	*40S ribosomal protein S12 (predicted)*	1	0	1	0	0	0	0	0	0	1	0	1
SPBC365.14c	***uge1***	*UDP-glucose 4-epimerase Uge1*	1	0	1	0	0	0	0	0	0	0	1	1
SPBC582.04c	***dsh1***	*RNAi protein, Dsh1*	0	1	1	0	0	0	0	0	0	1	1	0
SPCC1919.10c	***myo52***	*myosin type V*	0	1	1	0	0	0	0	0	0	1	1	0
SPBC11C11.07	***rpl1801***	*60S ribosomal protein L18*	0	1	1	0	0	0	0	0	0	0	1	1
SPCC553.01c	***dbl2***	*meiotic chromosome segregation protein Dbl2*	0	1	1	0	0	0	0	0	0	0	1	1
SPCC13B11.01	***adh1***	*alcohol dehydrogenase Adh1*	0	1	1	0	0	0	0	0	0	0	1	1
SPBP22H7.08	***rps1002***	*40S ribosomal protein S10 (predicted)*	0	1	1	0	0	0	0	0	0	0	1	1
SPCC126.09	***zip2***	*vacuolar membrane zinc transmembrane transporter (predicted)*	0	1	1	0	0	0	0	0	0	0	1	1
SPCC1494.07		*tRNA 2'-O-methylase subunit Trm72 (predicted)*	0	1	1	0	0	0	0	0	0	0	1	1
SPBC557.02c		*conserved fungal protein*	0	1	1	0	0	0	0	0	0	0	1	1
SPAC7D4.07c	***trx1***	*cytosolic thioredoxin Trx1*	0	1	1	0	0	0	0	0	0	0	1	1
SPCC1322.12c	***bub1***	*mitotic spindle checkpoint kinase Bub1*	0	1	1	0	0	0	0	0	0	0	1	1
SPBC3H7.09	***erf2***	*palmitoyltransferase Erf2*	0	1	0	0	0	0	0	0	0	1	1	1
SPBP16F5.03c	***tra1***	*SAGA complex phosphatidylinositol pseudokinase Tra1*	0	1	0	0	0	0	0	0	0	1	1	1
SPCC794.12c	***mae2***	*malic enzyme, malate dehydrogenase (oxaloacetate decarboxylating), Mae2*	0	1	0	0	0	0	0	0	0	1	1	1
SPBPB2B2.01		*amino acid transmembrane transporter (predicted)*	0	1	0	0	0	0	0	0	0	1	1	1
SPBP8B7.24c	***atg8***	*autophagy associated protein Atg8*	0	0	1	0	0	0	0	0	0	1	1	1
SPAC19G12.05		*mitochondrial citrate transmembrane transporter (predicted)*	0	0	1	0	0	0	0	0	0	1	1	1
SPAC4H3.03c		*glucan 1,4-alpha-glucosidase (predicted)*	0	0	1	0	0	0	0	0	0	1	1	1
SPAC1071.04c	***spc2***	*signal peptidase subunit Spc2 (predicted)*	0	0	1	0	0	0	0	0	0	1	1	1
SPAPB1E7.04c		*chitinase (predicted)*	0	0	0	0	0	0	1	1	1	0	0	0
SPAC1782.11	***met14***	*adenylyl-sulfate kinase (predicted)*	0	0	0	1	1	0	1	0	0	0	0	0
SPBC29A3.08	***pof4***	*elongin-A, F-box protein Pof4 (predicted)*	1	0	0	0	0	0	1	0	0	1	0	0
SPAC6G9.12	***cfr1***	*Chs five related protein Cfr1*	0	0	1	0	0	0	0	1	1	0	0	0
SPAC227.18	***lys3***	*saccharopine dehydrogenase Lys3*	0	0	0	0	0	0	0	1	1	0	1	0
SPAC688.11	***end4***	*Huntingtin-interacting protein homolog*	0	0	0	0	0	0	0	1	1	0	1	0
SPBC428.02c	***eca39***	*branched chain amino acid aminotransferase Eca39*	1	1	0	0	0	0	0	1	0	0	0	0
SPBC16C6.04	***dbl6***	*double strand break localizing protein Dbl6*	1	0	0	0	0	0	0	1	0	1	0	0
SPBC56F2.02	***rpl1901***	*60S ribosomal protein L19*	1	0	0	0	0	0	0	1	0	1	0	0
SPBC3B8.02	***php5***	*CCAAT-binding factor complex subunit Php5*	0	1	1	0	0	0	0	1	0	0	0	0
SPAPB1E7.14	***iec5***	*Ino80 complex subunit Iec5*	0	0	1	0	0	0	0	1	0	0	0	1
SPCC1281.03c	***emc4***	*ER membrane protein complex subunit 4 (predicted)*	0	0	0	0	0	0	0	1	0	0	1	1
SPAC22F8.12c	***shf1***	*small histone ubiquitination factor Shf1*	0	0	0	1	1	0	0	0	1	0	0	0
SPAC1B3.21	***coa3***	*cytochrome C oxidase assembly factor 3, Coa3 (predicted)*	0	1	0	0	1	0	0	0	1	0	0	0
SPAC3G9.03	***rpl2301***	*60S ribosomal protein L23*	0	0	0	0	0	1	0	0	1	1	0	0
SPBC1604.25	***pet117***	*cytochrome c oxidase assembly protein Pet117 (predicted)*	1	1	0	0	0	0	0	0	1	0	0	0
SPBC4C3.12	***sep1***	*forkhead transcription factor Sep1*	1	0	0	0	0	0	0	0	1	1	0	0
SPAC1002.15c	***med6***	*mediator complex subunit Med6*	1	0	0	0	0	0	0	0	1	1	0	0
SPCC188.02	***par1***	*protein phosphatase regulatory subunit Par1*	0	1	1	0	0	0	0	0	1	0	0	0
SPAC15E1.02c		*DUF1761 family protein*	0	1	0	0	0	0	0	0	1	0	1	0
SPAC18G6.05c	***gcn1***	*translation elongation regulator Gcn1 (predicted)*	0	0	0	0	0	0	0	0	1	1	1	0
SPCC1223.10c	***eaf1***	*RNA polymerase II transcription elongation factor SpEAF*	0	0	0	1	1	1	0	0	0	0	0	0
SPCC737.03c	***ima1***	*inner nuclear membrane protein Ima1*	0	0	0	1	1	1	0	0	0	0	0	0
SPCC1223.12c	***meu10***	*GPI anchored cell surface protein involved in ascospore wall assembly Meu10*	0	0	0	1	1	1	0	0	0	0	0	0
SPCC1223.06	***tea1***	*cell end marker Tea1*	0	0	0	1	1	1	0	0	0	0	0	0
SPCC1223.04c	***set11***	*ribosomal protein lysine methyltransferase Set11*	0	0	0	1	1	1	0	0	0	0	0	0
SPCC737.06c	***gcs2***	*glutamate-cysteine ligase regulatory subunit Gcs2 (predicted)*	0	0	0	1	1	1	0	0	0	0	0	0
SPAC9E9.09c	***atd1***	*aldehyde dehydrogenase (predicted)*	0	0	0	1	1	1	0	0	0	0	0	0

SPCC297.05		*diacylglycerol binding protein (predicted)*	0	0	0	1	1	0	0	0	0	0	1	0
SPCC1223.11	***ptc2***	*protein phosphatase 2C Ptc2*	0	0	0	1	1	0	0	0	0	0	1	0
SPBC1105.02c	***lys4***	*homocitrate synthase*	1	0	1	1	0	0	0	0	0	0	0	0
SPBC3B9.11c	***ctf1***	*mRNA cleavage and polyadenylation specificity factor complex subunit Ctf1*	1	0	0	1	0	0	0	0	0	1	0	0
SPCC31H12.05c	***sds21***	*serine/threonine protein phosphatase PP1 subfamily, Sds21*	1	0	0	1	0	0	0	0	0	1	0	0
SPBC27B12.08	***sip1***	*Pof6 interacting protein Sip1, predicted AP-1 accessory protein*	1	0	1	0	1	0	0	0	0	0	0	0
SPCC777.13	***vps35***	*retromer complex subunit Vps35*	1	0	0	0	1	0	0	0	0	1	0	0
SPBC713.11c	***pmp3***	*plasma membrane proteolipid Pmp3*	1	0	0	0	1	0	0	0	0	1	0	0
SPBC1734.08	***hse1***	*STAM like protein Hse1*	0	1	1	0	1	0	0	0	0	0	0	0
SPBC4F6.08c	***mrpl39***	*mitochondrial ribosomal protein subunit L39 (predicted)*	0	1	0	0	1	0	0	0	0	0	1	0
SPBC428.10		*Schizosaccharomyces pombe specific protein*	0	1	0	0	1	0	0	0	0	0	1	0
SPAC1F12.04c		*conserved fungal protein*	0	0	1	0	1	0	0	0	0	1	0	0
SPAPB1A11.01	***mfc1***	*copper transmembrane transporter, meiotic Mfc1*	0	0	1	0	1	0	0	0	0	0	1	0
SPAC13G7.06	***met16***	*phosphoadenosine phosphosulfate reductase*	0	0	0	0	1	0	0	0	0	1	1	0
SPCC645.14c	***sti1***	*chaperone activator Sti1 (predicted)*	0	0	0	0	1	0	0	0	0	1	1	0
SPAC144.02	***iec1***	*Ino80 complex subunit Iec1*	0	0	0	0	1	0	0	0	0	1	1	0
SPAC12G12.13c	***cid14***	*poly(A) polymerase Cid14*	1	0	0	0	0	1	0	0	0	1	0	0
SPBC3D6.04c	***mad1***	*mitotic spindle checkpoint protein Mad1*	1	0	0	0	0	1	0	0	0	1	0	0
SPAC19G12.12	***dlp1***	*decaprenyl diphosphate synthase subunit 2 Dlp1*	0	1	1	0	0	1	0	0	0	0	0	0
SPAC110.02	***pds5***	*mitotic cohesin-associated protein Pds5*	0	1	0	0	0	1	0	0	0	0	1	0
SPBC19F8.06c	***meu22***	*amino acid transmembrane transporter, predicted Meu22*	0	1	0	0	0	1	0	0	0	0	1	0
SPCC61.02	***spt3***	*SAGA complex subunit Spt3*	0	1	0	0	0	1	0	0	0	0	0	1
SPBC1861.01c	***cnp3***	*CENP-C ortholog Cnp3*	0	0	1	0	0	1	0	0	0	0	1	0
SPAC630.15	***mug177***	*Schizosaccharomyces pombe specific protein*	0	0	1	0	0	1	0	0	0	0	1	0
SPAC20H4.02	***dsc3***	*Golgi Dsc E3 ligase complex subunit Dsc3*	0	0	1	0	0	1	0	0	0	0	0	1
SPCC18.09c	***hnt3***	*aprataxin Hnt3*	0	0	1	0	0	1	0	0	0	0	0	1
SPCC74.06	***mak3***	*histidine kinase Mak3*	0	0	1	0	0	1	0	0	0	0	0	1
SPAC824.02	***bst1***	*GPI inositol deacylase Bst1 (predicted)*	0	0	0	0	0	1	0	0	0	1	1	0
SPAC750.08c		*NAD-dependent malic enzyme (predicted), partial*	0	0	0	0	0	1	0	0	0	1	1	0
SPBC405.07	***rpl3602***	*60S ribosomal protein L36*	0	0	0	0	0	1	0	0	0	1	1	0
SPBC29A3.12	***rps902***	*40S ribosomal protein S9 (predicted)*	0	0	0	0	0	1	0	0	0	1	1	0
SPAC8C9.04		*Schizosaccharomyces specific protein*	0	0	0	0	0	1	0	0	0	1	1	0
SPBC2F12.15c	***pfa3***	*palmitoyltransferase Pfa3 (predicted)*	0	0	0	0	0	1	0	0	0	1	1	0
SPBC106.19		*Schizosaccharomyces specific protein*	0	0	0	0	0	1	0	0	0	0	1	1
SPCC1450.16c	***ptl1***	*triacylglycerol lipase ptl1*	0	0	0	0	0	1	0	0	0	0	1	1
SPBC428.05c	***arg12***	*argininosuccinate synthase Arg12*	1	1	1	0	0	0	0	0	0	0	0	0
SPBP35G2.07	***ilv1***	*acetolactate synthase catalytic subunit*	1	1	1	0	0	0	0	0	0	0	0	0
SPBC25H2.18	***cox20***	*mitochondrial respiratory chain complex IV assembly protein Cox20 (predicted)*	1	1	1	0	0	0	0	0	0	0	0	0
SPAC2F7.03c	***pom1***	*DYRK family protein kinase Pom1*	1	1	0	0	0	0	0	0	0	1	0	0
SPAC4C5.04	***rad31***	*SUMO activating enzyme E1-type Rad31*	1	1	0	0	0	0	0	0	0	1	0	0
SPCC584.01c		*sulfite reductase NADPH flavoprotein subunit (predicted)*	1	1	0	0	0	0	0	0	0	1	0	0
SPBC8D2.17	***gmh4***	*alpha-1,2-galactosyltransferase (predicted)*	1	1	0	0	0	0	0	0	0	1	0	0
SPBC1271.14		*glutamate N-acetyltransferase (predicted)*	1	1	0	0	0	0	0	0	0	0	1	0
SPBC2D10.17	***clr1***	*SHREC complex subunit Clr1*	1	1	0	0	0	0	0	0	0	0	1	0
SPBC16E9.03c	***coa1***	*mitochondrial inner membrane protein involved in respiratory chain complex IV assembly Coa1 (predicted)*	1	1	0	0	0	0	0	0	0	0	1	0
SPBC887.15c	***sur2***	*sphingosine hydroxylase Sur2*	1	1	0	0	0	0	0	0	0	0	1	0
SPAC4G8.11c	***atp10***	*mitochondrial F1-F0 ATPase assembly protein (predicted)*	1	1	0	0	0	0	0	0	0	0	1	0
SPBC16E9.14c	***zrg17***	*cation diffusion family zinc transmembrane transporter Zrg17*	1	1	0	0	0	0	0	0	0	0	1	0
SPCC1393.13		*protein carboxyl methyltransferase, implicated in DNA damage response*	1	0	1	0	0	0	0	0	0	1	0	0

SPBPB10D8.07c		*transmembrane transporter (predicted)*	1	0	1	0	0	0	0	0	0	1	0	0
SPAC11D3.18c		*plasma membrane nicotinic acid transmembrane transporter (predicted)*	1	0	1	0	0	0	0	0	0	1	0	0
SPAC5H10.04		*NADPH dehydrogenase (predicted)*	1	0	1	0	0	0	0	0	0	1	0	0
SPAC637.09		*ribonuclease H70 (predicted)*	1	0	1	0	0	0	0	0	0	1	0	0
SPBC1271.12	***kes1***	*oxysterol binding protein (predicted)*	1	0	1	0	0	0	0	0	0	1	0	0
SPBC83.13		*mitochondrial tricarboxylic acid transmembrane transporter (predicted)*	1	0	1	0	0	0	0	0	0	1	0	0
SPAC8C9.19		*conserved fungal protein*	1	0	1	0	0	0	0	0	0	1	0	0
SPAC23A1.03	***apt1***	*adenine phosphoribosyltransferase (APRT) (predicted)*	1	0	1	0	0	0	0	0	0	1	0	0
SPBC14F5.09c	***ade8***	*adenylosuccinate lyase Ade8*	1	0	1	0	0	0	0	0	0	1	0	0
SPAC644.06c	***cdr1***	*NIM1 family serine/threonine protein kinase Cdr1/Nim1*	1	0	1	0	0	0	0	0	0	0	1	0
SPAC25A8.01c	***fft3***	*SMARCAD1 family ATP-dependent DNA helicase Fft3*	1	0	1	0	0	0	0	0	0	0	1	0
SPBC16D10.08c	***hsp104***	*heat shock protein Hsp104 (predicted)*	1	0	1	0	0	0	0	0	0	0	1	0
SPBP16F5.05c	***yar1***	*ribosome biogenesis protein Yar1 (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPCC1919.01	***ckk2***	*calmodulin-dependent kinase kinase 2*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC1F5.08c	***yam8***	*stretch-activated calcium ion channel Yam8*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC725.03		*pyridoxamine 5'-phosphate oxidase (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC19G12.17	***erh1***	*enhancer of rudimentary homolog Erh1*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC56F8.02		*AMP binding enzyme (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC27D7.03c	***mei2***	*RNA-binding protein involved in meiosis Mei2*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC6F12.03c	***fsv1***	*SNARE Fsv1*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC725.01		*aspartate aminotransferase (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPCC24B10.06		*Schizosaccharomyces specific protein, predicted GPI anchor*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC1556.03	***azr1***	*serine/threonine protein phosphatase Azr1*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC22F8.04	***pet1***	*phosphoenolpyruvate transmembrane transporter Pet1*	1	0	0	0	0	0	0	0	0	1	1	0
SPCC970.07c	***raf2***	*Rik1-associated factor Raf2*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC16A3.17c		*transmembrane transporter (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC354.03	***swd3***	*WD repeat protein Swd3*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC4.02c		*conserved fungal protein*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC1527.02	***sft2***	*Golgi transport protein Sft2 (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC27.08c	***sua1***	*sulfate adenylyltransferase*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC4G9.10	***arg3***	*ornithine carbamoyltransferase Arg3*	1	0	0	0	0	0	0	0	0	1	1	0
SPCC24B10.22	***pog1***	*mitochondrial DNA polymerase gamma*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC17G6.13	***slt1***	*Schizosaccharomyces specific protein Slt1*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC16E9.12c	***pab2***	*poly(A) binding protein Pab2*	1	0	0	0	0	0	0	0	0	1	1	0
SPCC11E10.07c		*translation initiation factor eIF2B alpha subunit (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC11G7.02	***pub1***	*HECT-type ubiquitin-protein ligase E3 Pub1*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC3H7.11	***abp140***	*actin binding methyltransferase Abp140 (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPCC330.03c		*NADPH-hemoprotein reductase (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPAC1556.04c	***cdd1***	*cytidine deaminase Ccd1 (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC16E9.13	***ksp1***	*serine/threonine protein kinase Ksp1 (predicted)*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC16E9.19		*conserved fungal protein*	1	0	0	0	0	0	0	0	0	1	1	0
SPBC32F12.08c	***duo1***	*DASH complex subunit Duo1*	1	1	0	0	0	0	0	0	0	1	1	0
SPBC1734.11	***mas5***	*DNAJ domain protein Mas5 (predicted)*	1	0	0	0	0	0	0	0	0	1	0	1
SPBC19G7.16	***iws1***	*transcription elongation factor complex subunit Iws1 (predicted)*	0	1	1	0	0	0	0	0	0	1	0	0
SPCC777.06c		*hydrolase (predicted)*	0	1	1	0	0	0	0	0	0	0	1	0
SPCC24B10.10c	***yta4***	*mitochondrial outer membrane ATPase Msp1/Yta4 (predicted)*	0	1	1	0	0	0	0	0	0	0	1	0
SPCC24B10.04		*Schizosaccharomyces specific protein*	0	1	1	0	0	0	0	0	0	0	1	0
SPBC24C6.11	***cwf14***	*G10 protein*	0	1	1	0	0	0	0	0	0	0	1	0
SPBC6B1.09c	***nbs1***	*Mre11 complex subunit Nbs1*	0	1	1	0	0	0	0	0	0	0	1	0

SPAC30.01c	***sec72***	*Sec7 domain protein, ARF GEF Sec72*	0	1	1	0	0	0	0	0	0	0	1	0
SPAC1B1.04c	***pan3***	*protein kinase-like PAN complex subunit Pan3 (predicted)*	0	1	1	0	0	0	0	0	0	0	1	0
SPAC31G5.19	***abo1***	*ATPase with bromodomain protein (predicted)*	0	1	1	0	0	0	0	0	0	0	1	0
SPBC1709.09	***rrf1***	*mitochondrial translation termination factor Rrf1*	0	1	1	0	0	0	0	0	0	0	1	0
SPCC594.02c		*conserved fungal protein*	0	1	1	0	0	0	0	0	0	0	1	0
SPCC126.04c	***sgf73***	*SAGA complex subunit Sgf73*	0	1	1	0	0	0	0	0	0	0	1	0
SPCC63.14	***eis1***	*eisosome assembly protein eis1*	0	1	1	0	0	0	0	0	0	0	1	0
SPAC1F3.07c	***rsc58***	*RSC complex subunit Rsc58*	0	1	1	0	0	0	0	0	0	0	1	0
SPCC645.07	***rgf1***	*RhoGEF for Rho1, Rgf1*	0	1	1	0	0	0	0	0	0	0	1	0
SPAC1F7.07c	***fip1***	*iron permease Fip1*	0	1	1	0	0	0	0	0	0	0	1	0
SPBC557.04	***ppk29***	*Ark1/Prk1 family protein kinase Ppk29*	0	1	1	0	0	0	0	0	0	0	0	1
SPCP1E11.11	***puf6***	*Puf family RNA-binding protein Puf6 (predicted)*	0	1	1	0	0	0	0	0	0	0	0	1
SPCC663.03	***pmd1***	*leptomycin transmembrane transporter Pmd1*	0	1	1	0	0	0	0	0	0	0	0	1
SPAC25B8.03	***psd2***	*phosphatidylserine decarboxylase Psd2*	0	1	1	0	0	0	0	0	0	0	0	1
SPBC17A3.09c	***aim22***	*lipoate-protein ligase A (predicted)*	0	1	1	0	0	0	0	0	0	0	0	1
SPAC22E12.18		*conserved fungal protein*	0	1	0	0	0	0	0	0	0	1	1	0
SPAC24B11.09	***mpc2***	*mitochondrial pyruvate transmembrane transporter Mpc2 (predicted)*	0	1	0	0	0	0	0	0	0	1	1	0
SPCC1393.12		*Schizosaccharomyces specific protein*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC3H7.18	***tam8***	*Schizosaccharomyces specific protein Tam8*	0	1	0	0	0	0	0	0	0	1	1	0
SPAC15A10.15	***sgo2***	*inner centromere protein, shugoshin Sgo2*	0	1	0	0	0	0	0	0	0	1	1	0
SPAC17H9.04c	***nrp1***	*RNA-binding protein*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC27B12.07		*conserved fungal protein*	0	1	0	0	0	0	0	0	0	1	1	0
SPBPB10D8.06c		*transmembrane transporter (predicted)*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC776.06c		*Arf3/6 docking factor (predicted)*	0	1	0	0	0	0	0	0	0	1	1	0
SPCC31H12.08c	***ccr4***	*CCR4-Not complex subunit Ccr4 (predicted)*	0	1	0	0	0	0	0	0	0	1	1	0
SPBP16F5.08c		*flavin dependent monooxygenase (predicted)*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC9B6.03		*zf-FYVE type zinc finger protein*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC3H7.10	***elp6***	*elongator complex subunit Elp6 (predicted)*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC2F12.12c	***cay1***	*cactin, spliceosome complex subunit*	0	1	0	0	0	0	0	0	0	1	1	0
SPAC10F6.08c	***nht1***	*Ino80 complex HMG box subunit Nht1*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC83.10		*transthyretin superfamily member, human ER membrane protein complex subunit 7 ortholog*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC16H5.06	***rip1***	*ubiquinol-cytochrome-c reductase complex subunit 5*	0	1	0	0	0	0	0	0	0	1	1	0
SPAC17C9.08	***pnu1***	*mitochondrial endodeoxyribonuclease Pnu1*	0	1	0	0	0	0	0	0	0	1	1	0
SPAC13A11.05	***ysp2***	*peptidase family M17 cytoplasmic leucyl aminopeptidase yspII (LAP yspII)*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC30D10.13c	***pdb1***	*pyruvate dehydrogenase e1 component beta subunit Pdb1*	0	1	0	0	0	0	0	0	0	1	1	0
SPCC74.09	***mug24***	*RNA-binding protein, rrm type*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC17D1.02	***dph2***	*diphthamide biosynthesis protein (predicted)*	0	1	0	0	0	0	0	0	0	1	1	0
SPBC16E9.06c	***uvi31***	*BolA domain UV induced protein Uvi31*	0	1	0	0	0	0	0	0	0	1	1	0
SPCC306.04c	***set1***	*histone lysine methyltransferase Set1*	0	1	0	0	0	0	0	0	0	1	1	0
SPCC191.11	***inv1***	*external invertase, beta-fructofuranosidase*	0	1	0	0	0	0	0	0	0	0	1	1
SPCC31H12.04c	***rpl1202***	*60S ribosomal protein L12.1/L12A*	0	1	0	0	0	0	0	0	0	0	1	1
SPBP35G2.08c	***air1***	*zinc knuckle TRAMP complex subunit Air1*	0	1	0	0	0	0	0	0	0	0	1	1
SPCC1902.02	***mug72***	*oxidoreductase (predicted)*	0	1	0	0	0	0	0	0	0	0	1	1
SPCC663.06c	***osr1***	*short chain dehydrogenase (predicted)*	0	1	0	0	0	0	0	0	0	0	1	1
SPBC56F2.08c		*RNA-binding protein (predicted)*	0	1	0	0	0	0	0	0	0	0	1	1
SPBC9B6.11c		*CCR4/nocturin family endoribonuclease (predicted)*	0	0	1	0	0	0	0	0	0	1	1	0
SPBC3H7.12	***rav2***	*RAVE complex subunit Rav2*	0	0	1	0	0	0	0	0	0	1	1	0
SPAC664.02c	***arp8***	*actin-like protein, Ino80 complex subunit Arp8*	0	0	1	0	0	0	0	0	0	1	1	0

SPAC328.09		mitochondrial 2-oxoadipate and 2-oxoglutarate transmembrane transporter (predicted)	0	0	1	0	0	0	0	0	0	1	1	0
SPAC22E12.11c	**set3**	histone lysine methyltransferase Set3	0	0	1	0	0	0	0	0	0	1	1	0
SPAC1B3.16c	**vht1**	vitamin H transmembrane transporter Vht1	0	0	1	0	0	0	0	0	0	1	1	0
SPBC1703.12	**ubp9**	ubiquitin C-terminal hydrolase Ubp9	0	0	1	0	0	0	0	0	0	1	1	0
SPCC1672.11c		P-type ATPase P5 type (predicted)	0	0	1	0	0	0	0	0	0	1	1	0
SPBC1198.11c	**reb1**	RNA polymerase I transcription termination factor/ RNA polymerase II transcription factor Reb1	0	0	1	0	0	0	0	0	0	1	1	0
SPBC1348.02		S. pombe specific 5Tm protein family	0	0	1	0	0	0	0	0	0	1	1	0
SPAC20G8.09c	**nat10**	ribosome biogenesis ATPase	0	0	1	0	0	0	0	0	0	1	1	0
SPCC613.12c	**raf1**	CLRC ubiquitin E3 ligase complex specificiy factor Raf1/Dos1	0	0	1	0	0	0	0	0	0	1	0	1
SPBP8B7.13	**vac7**	Vac7 ortholog (predicted)	0	0	1	0	0	0	0	0	0	1	0	1
SPCC285.04		transthyretin/hydroxyisourate hydrolase (predicted)	0	0	1	0	0	0	0	0	0	1	0	1
SPCC622.18	**rpl6**	60S ribosomal protein L6 (predicted)	0	0	1	0	0	0	0	0	0	1	0	1
SPAC26F1.01	**sec74**	guanyl-nucleotide exchange factor Sec74 (predicted)	0	0	1	0	0	0	0	0	0	1	0	1
SPAC3C7.12	**tip1**	CLIP170 family protein Tip1	0	0	1	0	0	0	0	0	0	0	1	1
SPBC21B10.02		conserved fungal protein	0	0	1	0	0	0	0	0	0	0	1	1
SPBC776.03		homoserine dehydrogenase (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPAC23H4.17c	**srb10**	cyclin-dependent protein Srb mediator subunit kinase Srb10	0	0	1	0	0	0	0	0	0	0	1	1
SPBC337.13c	**gtr1**	Gtr1/RagA G protein Gtr1 (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPBC409.18		phosphatidic acid phosphatase (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPAC3A11.06	**mvp1**	sorting nexin Mvp1 (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPBC21C3.01c	**vps1301**	chorein homolog Vps13a (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPBC29A3.21		Schizosaccharomyces pombe specific protein	0	0	1	0	0	0	0	0	0	0	1	1
SPCC569.06		S. pombe specific multicopy membrane protein family 1	0	0	1	0	0	0	0	0	0	0	1	1
SPCC1322.02	**pxd1**	structure-specific DNA nuclease regulator Pxd1	0	0	1	0	0	0	0	0	0	0	1	1
SPBC83.09c	**lin1**	U5 snRNP subunit Snu40 (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPBPJ4664.05		conserved fungal protein	0	0	1	0	0	0	0	0	0	0	1	1
SPAC4F8.03	**sdo1**	SBDS family ribosome maturation protein Sdo1 (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPCC962.05	**ast1**	asteroid homolog, XP-G family protein	0	0	1	0	0	0	0	0	0	0	1	1
SPAC5D6.09c	**mug86**	acetate transmembrane transporter (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPCC1259.12c	**gid1**	GID complex subunit, Ran GTPase binding protein Gid1 (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPCC132.04c	**gdh2**	NAD-dependent glutamate dehydrogenase Gdh2 (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPCC306.11		Schizosaccharomyces specific protein	0	0	1	0	0	0	0	0	0	0	1	1
SPCC306.08c	**mdh1**	malate dehydrogenase Mdh1 (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPCC1442.03	**mme1**	mitochondrial magnesium ion transmembrane transporter Mme1 (predicted)	0	0	1	0	0	0	0	0	0	0	1	1
SPCC4B3.15	**mid1**	medial ring protein Mid1	0	0	1	0	0	0	0	0	0	0	1	1
SPBP23A10.05	**ssr4**	SWI/SNF and RSC complex subunit Ssr4	0	0	0	0	0	0	0	0	0	1	1	1
SPAC18B11.11		GTPase activating protein (predicted)	0	0	0	0	0	0	0	0	0	1	1	1
SPBC21C3.07c	**trm140**	tRNA (cytosine) methyltransferase Trm140 (predicted)	0	0	0	0	0	0	1	1	0	0	0	0
SPAC1486.10	**thi1**	transcription factor Thi1	0	0	0	0	0	0	1	1	0	0	0	0
SPAC1486.01		manganese superoxide dismutase	0	0	0	0	0	0	1	0	1	0	0	0
SPBC26H8.03	**cho2**	phosphatidylethanolamine N-methyltransferase Cho2	0	0	0	0	0	1	1	0	0	0	0	0
SPAPB1E7.05	**gde1**	glycerophosphoryl diester phosphodiesterase Gde1 (predicted)	0	0	0	0	0	0	0	1	1	0	0	0
SPAC959.08	**rpl2102**	60S ribosomal protein L21 (predicted)	0	0	0	0	0	0	0	1	1	0	0	0
SPAC6G9.16c	**xrc4**	XRCC4 nonhomologous end joining factor Xrc4	0	0	0	0	0	0	0	1	1	0	0	0
SPAC6G9.09c	**rpl2401**	60S ribosomal protein L24 (predicted)	0	0	0	0	0	1	0	1	0	0	0	0
SPCC4B3.14	**cwf20**	complexed with Cdc5 protein Cwf20	1	0	0	0	0	0	0	1	0	0	0	0

SPAC26A3.01	***sxa1***	*aspartic protease Sxa1*	0	0	1	0	0	0	0	1	0	0	0	0
SPAC3A11.10c		*dipeptidyl peptidase (predicted)*	0	0	1	0	0	0	0	1	0	0	0	0
SPBC56F2.12	***ilv5***	*acetohydroxyacid reductoisomerase (predicted)*	0	1	0	0	0	0	0	0	1	0	0	0
SPAC26A3.02	***myh1***	*adenine DNA glycosylase Myh1*	0	0	1	0	0	0	0	0	1	0	0	0
SPAC6G10.08	***idp1***	*isocitrate dehydrogenase Idp1 (predicted)*	0	0	1	0	0	0	0	0	1	0	0	0
SPAC26A3.09c	***rga2***	*Rho-type GTPase activating protein Rga2*	0	0	1	0	0	0	0	0	1	0	0	0
SPAC31G5.12c	***maf1***	*repressor of RNA polymerase III Maf1*	0	0	1	0	0	0	0	0	1	0	0	0
SPAC6G9.03c	***mug183***	*histone chaperone Rtt106-like (predicted)*	0	0	0	0	0	0	0	0	1	0	1	0
SPBC12C2.01c		*Schizosaccharomyces specific protein*	0	0	0	0	0	0	0	0	1	0	1	0
SPAC12G12.17		*non-classical export protein 1 (predicted)*	0	0	0	0	0	0	0	0	1	0	1	0
SPBC2D10.11c	***nap2***	*nucleosome assembly protein Nap2*	0	0	0	0	0	0	0	0	1	0	1	0
SPAC2E1P5.02c	***mug109***	*Rab GTPase binding protein upregulated in meiosis II (predicted)*	0	0	0	0	0	0	0	0	1	0	0	1
SPBC31F10.14c	***hip3***	*HIRA interacting protein Hip3*	0	0	0	1	0	1	0	0	0	0	0	0
SPCC737.04		*UPF0300 family protein 6*	0	0	0	1	0	1	0	0	0	0	0	0
SPCC737.05		*peroxin Pex28/29 (predicted)*	0	0	0	1	0	1	0	0	0	0	0	0
SPCC737.07c		*DNA polymerase alpha-associated DNA helicase A (predicted)*	0	0	0	1	0	1	0	0	0	0	0	0
SPCC1223.03c	***gut2***	*glycerol-3-phosphate dehydrogenase Gut2 (predicted)*	0	0	0	1	0	1	0	0	0	0	0	0
SPCC162.05	***coq3***	*hexaprenyldihydroxybenzoate methyltransferase Coq3*	1	0	0	1	0	0	0	0	0	0	0	0
SPCC285.05		*purine nucleoside transmembrane transporter (predicted)*	0	1	0	1	0	0	0	0	0	0	0	0
SPCC1840.03	***sal3***	*karyopherin Sal3*	0	0	0	0	1	1	0	0	0	0	0	0
SPAC22F8.05		*alpha,alpha-trehalose-phosphate synthase (predicted)*	0	0	0	0	1	1	0	0	0	0	0	0
SPAC7D4.03c		*conserved fungal family*	0	0	0	0	1	1	0	0	0	0	0	0
SPCC736.06	***dar2***	*mitochondrial aspartate-tRNA ligase Dar2 (predicted)*	1	0	0	0	1	0	0	0	0	0	0	0
SPAC4A8.03c	***ptc4***	*protein phosphatase 2C Ptc4*	1	0	0	0	1	0	0	0	0	0	0	0
SPAC1A6.04c	***plb1***	*phospholipase B homolog Plb1*	1	0	0	0	1	0	0	0	0	0	0	0
SPAC27E2.03c		*Obg-Like ATPase (predicted)*	0	0	1	0	1	0	0	0	0	0	0	0
SPCC1840.11	***csl4***	*exosome subunit Csl4*	0	0	0	0	1	0	0	0	0	0	1	0
SPAC24B11.06c	***sty1***	*MAP kinase Sty1*	0	0	0	0	1	0	0	0	0	0	1	0
SPCC1450.06c	***grx3***	*monothiol glutaredoxin Grx3*	0	0	0	0	1	0	0	0	0	0	0	1
SPCC338.10c	***cox5***	*cytochrome c oxidase subunit V (predicted)*	1	0	0	0	0	1	0	0	0	0	0	0
SPBC337.16	***cho1***	*phosphatidyl-N-dimethylethanolamine N-methyltransferase*	1	0	0	0	0	1	0	0	0	0	0	0
SPAC1687.12c	***coq4***	*ubiquinone biosynthesis protein Coq4 (predicted)*	0	1	0	0	0	1	0	0	0	0	0	0
SPBC30B4.03c	***adn1***	*adhesion defective protein*	0	1	0	0	0	1	0	0	0	0	0	0
SPAC589.11	***pth4***	*mitochondrial translation release factor*	0	0	1	0	0	1	0	0	0	0	0	0
SPBC4F6.15c	***swi10***	*DNA repair endonuclease non-catalytic subunit Swi10*	0	0	1	0	0	1	0	0	0	0	0	0
SPCC162.06c		*vacuolar sorting protein Vps60 (predicted)*	0	0	1	0	0	1	0	0	0	0	0	0
SPAC8E11.01c		*beta-fructofuranosidase (predicted)*	0	0	1	0	0	1	0	0	0	0	0	0
SPBC660.07	***ntp1***	*alpha,alpha-trehalase Ntp1*	0	0	0	0	0	1	0	0	0	1	0	0
SPAC13A11.03	***mcp7***	*meiosis specific coiled-coil protein Mcp7*	0	0	0	0	0	1	0	0	0	1	0	0
SPBC2D10.19c	***alb1***	*pre-60S shuttling factor Alb1 (predicted)*	0	0	0	0	0	1	0	0	0	0	1	0
SPBC1709.12	***rid1***	*GTPase binding protein Rid1 (predicted)*	0	0	0	0	0	1	0	0	0	0	1	0
SPBP4H10.05c	***spe2***	*S-adenosylmethionine decarboxylase proenzyme Spe2*	0	0	0	0	0	1	0	0	0	0	1	0
SPCC18.11c	***sdc1***	*Dpy-30 domain protein Sdc1*	0	0	0	0	0	1	0	0	0	0	1	0
SPAC6F6.06c	***rax2***	*cell polarity factor Rax2*	0	0	0	0	0	1	0	0	0	0	1	0
SPAC1296.01c		*phosphoacetylglucosamine mutase (predicted)*	0	0	0	0	0	1	0	0	0	0	1	0
SPAC694.04c		*conserved eukaryotic protein*	0	0	0	0	0	1	0	0	0	0	1	0
SPAC222.08c	***sno1***	*glutamine aminotransferase subunit Sno1 (predicted)*	0	0	0	0	0	1	0	0	0	0	1	0
SPAC31G5.18c	***sde2***	*silencing defective protein Sde2*	0	0	0	0	0	1	0	0	0	0	1	0
SPCC1620.02	***wtf23***	*wtf element Wtf23*	0	0	0	0	0	1	0	0	0	0	0	1
SPAC1F5.10		*ATP-dependent RNA helicase (predicted)*	1	1	0	0	0	0	0	0	0	0	0	0
SPCC794.07	***lat1***	*dihydrolipoamide S-acetyltransferase E2, Lat1 (predicted)*	1	1	0	0	0	0	0	0	0	0	0	0

SPAC821.11	***pro1***	*gamma-glutamyl phosphate reductase Pro1 (predicted)*	1	1	0	0	0	0	0	0	0	0	0	0
SPBC83.18c	***fic1***	*C2 domain protein Fic1*	1	1	0	0	0	0	0	0	0	0	0	0
SPBC19F8.08	***rps401***	*40S ribosomal protein S4 (predicted)*	1	1	0	0	0	0	0	0	0	0	0	0
SPAC8C9.06c	***ppr4***	*mitochondrial translation regulator Ppr4*	1	1	0	0	0	0	0	0	0	0	0	0
SPAC56F8.04c	***ppt1***	*para-hydroxybenzoate--polyprenyltransferase Ppt1*	1	1	0	0	0	0	0	0	0	0	0	0
SPBC16A3.03c	***ppr7***	*mitochondrial PPR repeat protein Ppr7*	1	1	0	0	0	0	0	0	0	0	0	0
SPCC16C4.13c	***rpl1201***	*60S ribosomal protein L12.1/L12A*	1	1	0	0	0	0	0	0	0	0	0	0
SPBC337.08c	***ubi4***	*ubiquitin*	1	0	1	0	0	0	0	0	0	0	0	0
SPAC17G8.06c		*dihydroxy-acid dehydratase (predicted)*	1	0	1	0	0	0	0	0	0	0	0	0
SPBC146.12	***coq6***	*monooxygenase Coq6 (predicted)*	1	0	1	0	0	0	0	0	0	0	0	0
SPAC19G12.11	***coq9***	*ubiquinone biosynthesis protein Coq9 (predicted)*	1	0	1	0	0	0	0	0	0	0	0	0
SPCC320.12	***atp23***	*mitochondrial inner membrane peptidase Atp23 (predicted)*	1	0	1	0	0	0	0	0	0	0	0	0
SPAC19E9.03	***pas1***	*cyclin Pas1*	1	0	1	0	0	0	0	0	0	0	0	0
SPBC31F10.10c		*zf-MYND type zinc finger protein*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC1783.08c	***rpl1502***	*60S ribosomal protein L15b (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC18B5.10c		*TREX complex subunit Tex1 (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPBC21.02	***rtc5***	*TLDc domain protein 2*	1	0	0	0	0	0	0	0	0	1	0	0
SPBC660.11	***tcg1***	*single-stranded telomeric binding protein Tgc1*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC17G8.05	***med20***	*mediator complex subunit Med20*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC1827.02c	***pcy1***	*cholinephosphate cytidylyltransferase Pcy1 (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC630.07c		*Schizosaccharomyces specific protein*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC29E6.01	***pof11***	*F-box protein Pof11*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC1020.06c	***tal1***	*transaldolase (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPBC16A3.16	***coa5***	*mitochondrial inner membrane protein involved in cytochrome c oxidase assembly Coa5 (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC513.03	***mfm2***	*M-factor precursor Mfm2*	1	0	0	0	0	0	0	0	0	1	0	0
SPBC16A3.06	***tad1***	*tRNA specific adenosine-37 deaminase Tad1 (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC1393.08		*transcription factor, zf-GATA type (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC637.03		*conserved fungal protein*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC338.05c	***mms2***	*ubiquitin conjugating enzyme Mms2*	1	0	0	0	0	0	0	0	0	1	0	0
SPBC609.02	***ptn1***	*phosphatidylinositol-3,4,5-trisphosphate3-phosphatase Ptn1*	1	0	0	0	0	0	0	0	0	1	0	0
SPBC16C6.08c	***qcr6***	*ubiquinol-cytochrome-c reductase complex subunit 8, hinge protein (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC1296.04	***mug65***	*spore wall assembly protein Mug65 (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC1851.03	***ckb1***	*CK2 family regulatory subunit Ckb1*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC162.11c		*uridine kinase/uracil phosphoribosyltransferase (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC790.02	***pep3***	*HOPS/CORVET complex subunit, ubiquitin-protein ligase E3 (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC977.06		*S. pombe specific DUF999 family protein 3*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC622.15c		*Schizosaccharomyces specific protein*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC26H5.10c	***tif51***	*translation elongation factor eIF5A (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC22F3.08c	***rok1***	*ATP-dependent RNA helicase Rok1 (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC25B8.13c	***isp7***	*2-OG-Fe(II) oxygenase superfamily protein*	1	0	0	0	0	0	0	0	0	1	0	0
SPBC3E7.09		*Sad1-UNC-like protein involved protein folding in the ER (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC29B12.05c		*mitochondrial S-adenosylmethionine-dependent methyltransferase (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPBC1271.05c		*zf-AN1 type zinc finger protein*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC18B5.06	***dom34***	*peloto ortholog (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC5D6.04		*auxin family transmembrane transporter (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC20G4.07c	***sts1***	*C-24(28) sterol reductase Sts1*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC27E2.01		*alpha-amylase homolog (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0

SPAC27D7.09c		*But2 family protein*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC1039.08		*serine acetyltransferase (predicted)*	1	0	0	0	0	0	0	0	0	1	0	0
SPCC188.07	***ccq1***	*telomere maintenance protein Ccq1*	1	0	0	0	0	0	0	0	0	1	0	0
SPAC13C5.02	***dre4***	*splicing associated factor Dre4*	1	0	0	0	0	0	0	0	0	0	1	0
SPBC106.04	***ada1***	*adenosine deaminase Ada1*	1	0	0	0	0	0	0	0	0	0	1	0
SPBC16C6.03c		*ribosome assembly protein (predicted)*	1	0	0	0	0	0	0	0	0	0	1	0
SPAC24H6.07	***rps901***	*40S ribosomal protein S9*	1	0	0	0	0	0	0	0	0	0	1	0
SPAC17A2.06c	***vps8***	*WD repeat protein Vps8 (predicted)*	1	0	0	0	0	0	0	0	0	0	1	0
SPAC630.10	***bmt2***	*rRNA (adenine) methyltransferase activity Bmt2 (predicted)*	1	0	0	0	0	0	0	0	0	0	1	0
SPBC15D4.03	***slm9***	*hira protein Slm9*	1	0	0	0	0	0	0	0	0	0	1	0
SPCC1739.07	***cti1***	*Cut3 interacting protein Cti1, predicted exosome subunit*	0	1	1	0	0	0	0	0	0	0	0	0
SPBC18H10.16	***can1***	*arginine transmembrane transporter Can1*	0	1	0	0	0	0	0	0	0	1	0	0
SPBC27.02c	***ask1***	*DASH complex subunit Ask1*	0	1	0	0	0	0	0	0	0	1	0	0
SPCC16A11.10c	***oca8***	*cytochrome b5 (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPCC550.07		*acetamidase (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC4G8.13c	***prz1***	*calcineurin responsive transcription factor Prz1*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC17A5.09c	***glc8***	*protein phosphatase regulatory subunit Glc8 (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC32H8.07	***git5***	*heterotrimeric G protein beta subunit Git5*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC1F7.09c		*allantoicase (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC12B10.09	***pet801***	*mitochondrial S-adenosylmethionine transmembrane transporter (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPCC24B10.02c		*NAD/NADH kinase (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC336.05c		*small RNA 2'-O-methyltransferase activity (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC4G9.16c	***rpl901***	*60S ribosomal protein L9*	0	1	0	0	0	0	0	0	0	0	1	0
SPCC830.10	***ham1***	*nucleoside triphosphatase Ham1 (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC4F6.12	***pxl1***	*paxillin-like protein Pxl1*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC1685.09	***rps29***	*40S ribosomal protein S29 (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC8E4.04		*alditol NADP+ 1-oxidoreductase activity (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC4D7.10c	***spt20***	*SAGA complex subunit Spt20*	0	1	0	0	0	0	0	0	0	0	1	0
SPBP8B7.22	***erd2***	*HDEL receptor (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC887.05c	***cwf29***	*RNA-binding protein Cwf29*	0	1	0	0	0	0	0	0	0	0	1	0
SPACUNK4.12c	***iph1***	*insulinase pombe homologue 1*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC725.02	***mpr1***	*histidine-containing response regulator phosphotransferase Mpr1*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC649.04	***uvi15***	*UV-induced protein Uvi15*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC3B9.13c	***rpp102***	*60S acidic ribosomal protein A3*	0	1	0	0	0	0	0	0	0	0	1	0
SPCC1442.05c	***mic26***	*MICOS complex subunit Mic26 (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC15E1.04	***hal3***	*thymidylate synthase/ flavoprotein fusion protein Hal3*	0	1	0	0	0	0	0	0	0	0	1	0
SPCC338.16	***pof3***	*F-box protein Pof3*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC16C6.02c	***vps1302***	*chorein homolog Vps1302 (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPCC1442.13c		*RNA-binding protein, G-patch type*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC9G1.07		*Schizosaccharomyces specific protein*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC21B10.15		*Schizosaccharomyces specific protein*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC1604.03c		*conserved fungal protein*	0	1	0	0	0	0	0	0	0	0	1	0
SPCC553.12c		*transmembrane transporter (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPAPB2B4.06		*conserved fungal protein*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC14C8.09c	***dbl3***	*IMPACT domain protein, possible chaperone (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC4D7.06c	***met8***	*siroheme synthase Met8 (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC6B1.04	***mde4***	*microtubule-site clamp monopolin complex subunit Mde4*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC6F6.19		*RNA-binding protein, G-patch type*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC1039.07c		*aminotransferase class-III, possible transaminase, unknown specificity*	0	1	0	0	0	0	0	0	0	0	1	0
SPBC460.02c		*eukaryotic translation elongation factor, glutathione S-transferase (predicted)*	0	1	0	0	0	0	0	0	0	0	1	0
SPAC4F10.04	***ypa1***	*protein phosphatase type 2A regulator, PTPA family Ypa1*	0	1	0	0	0	0	0	0	0	0	1	0

SPBC1306.02	**rtt10**	WD repeat protein, human WDR6 family, involved in endocytic recycling	0	1	0	0	0	0	0	0	0	0	1	0
SPBC14F5.13c	**pho8**	vacuolar membrane alkaline phosphatase (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPCC417.16		cytochrome c oxidase subunit (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPAC25H1.05	**meu29**	calcium transport regulatory factor (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPBC1718.03	**ker1**	DNA-directed RNA polymerase I complex subunit Ker1	0	1	0	0	0	0	0	0	0	0	1	0
SPCC1235.14	**ght5**	hexose transmembrane transporter Ght5	0	1	0	0	0	0	0	0	0	0	1	0
SPBC651.10	**nse5**	Smc5-6 complex non-SMC subunit Nse5	0	1	0	0	0	0	0	0	0	0	1	0
SPBC1778.02	**rap1**	telomere binding protein Rap1	0	1	0	0	0	0	0	0	0	0	1	0
SPAC644.15	**rpp101**	60S acidic ribosomal protein A1	0	1	0	0	0	0	0	0	0	0	1	0
SPAC31A2.09c	**apm4**	AP-2 adaptor complex subunit Apm4 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPAC11E3.08c	**nse6**	Smc5-6 complex non-SMC subunit Nse6	0	1	0	0	0	0	0	0	0	0	1	0
SPCC736.14	**dis1**	TOG/XMAP14 microtubule-associated protein Dis1 : defects -> spindle assembly checkpoint activation and mitotic arrest	0	1	0	0	0	0	0	0	0	0	1	0
SPCP1E11.10		ankyrin repeat protein, unknown biological role	0	1	0	0	0	0	0	0	0	0	1	0
SPAC13G7.02c	**ssa1**	heat shock protein Ssa1 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPAC17A5.07c	**ulp2**	SUMO deconjugating cysteine peptidase Ulp2 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPAC1851.04c	**ric1**	Ypt/Rab-specific guanyl-nucleotide exchange factor (GEF) subunit Ric1	0	1	0	0	0	0	0	0	0	0	1	0
SPAC29B12.06c	**rcd1**	RNA-binding protein, CCR4-NOT complex subunit Rcd1	0	1	0	0	0	0	0	0	0	0	1	0
SPAC3H1.06c		transmembrane transporter (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPBC25H2.15		SSU-rRNA maturation protein Tsr4 homolog 1 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPBC685.06	**rps001**	40S ribosomal protein S0A (p40)	0	1	0	0	0	0	0	0	0	0	1	0
SPCC622.12c	**gdh1**	NADP-specific glutamate dehydrogenase Gdh1 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPAC8F11.10c	**pvg1**	pyruvyltransferase Pvg1	0	1	0	0	0	0	0	0	0	0	1	0
SPAC1610.01	**saf5**	splicing factor Saf5	0	1	0	0	0	0	0	0	0	0	1	0
SPBC1778.06c	**fim1**	fimbrin	0	1	0	0	0	0	0	0	0	0	1	0
SPAC13G7.03	**upf3**	up-frameshift suppressor 3 family protein (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPAC14C4.06c	**nab2**	poly(A) binding protein Nab2 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPBC1778.03c		NADH pyrophosphatase (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPCC61.05		S. pombe specific multicopy membrane protein family 1	0	1	0	0	0	0	0	0	0	0	1	0
SPCC576.13	**swc5**	Swr1 complex subunit Swc5	0	1	0	0	0	0	0	0	0	0	1	0
SPACUNK12.02c	**cmk1**	calcium/calmodulin-dependent protein kinase Cmk1	0	1	0	0	0	0	0	0	0	0	1	0
SPAC1F7.12	**yak3**	aldose reductase ARK13 family YakC	0	1	0	0	0	0	0	0	0	0	1	0
SPAC25G10.09c	**pan1**	actin cortical patch component, with EF hand and WH2 motif Panl (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPCC63.06		human WDR89 family WD repeat protein	0	1	0	0	0	0	0	0	0	0	1	0
SPBC1604.20c	**tea2**	kinesin-like protein Tea2	0	1	0	0	0	0	0	0	0	0	1	0
SPCC4F11.02	**ptc1**	protein phosphatase 2C Ptc1	0	1	0	0	0	0	0	0	0	0	1	0
SPCC1183.02		glutathione S-transferase, translational elongation factor eEF1 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPCC364.05	**vps3**	CORVET complex subunit, GTPase regulator Vps3 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPCC191.01		Schizosaccharomyces specific protein	0	1	0	0	0	0	0	0	0	0	1	0
SPCC895.09c	**ucp12**	ATP-dependent RNA helicase Ucp12 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPAC1039.01		amino acid permease (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPAC17A5.08	**erp2**	COPII-coated vesicle component Erp2/3/4 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPCC1840.12	**opt3**	OPT oligopeptide transmembrane transporter family protein Opt3	0	1	0	0	0	0	0	0	0	0	1	0
SPCC569.04		Schizosaccharomyces pombe specific protein	0	1	0	0	0	0	0	0	0	0	1	0
SPBC660.10		mitochondrial translation elongation factor G (predicted)	0	1	0	0	0	0	0	0	0	0	1	0
SPBC839.03c		neddylation protein Dcn1 (predicted)	0	1	0	0	0	0	0	0	0	0	1	0

SPAC3A12.09c	***ure4***	*urease accessory protein UreD (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPCC14G10.03c	***ump1***	*proteasome maturation factor Ump1 (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPAC1F5.07c	***hem14***	*protoporphyrinogen oxidase Hem14 (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPCC338.08	***ctp1***	*CtIP-related endonuclease*	0	1	0	0	0	0	0	0	0	1	0
SPCC126.12		*GTP cyclohydrolase (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPAC23H3.09c	***gly1***	*threonine aldolase Gly1 (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPAC1142.07c	***vps32***	*ESCRT III complex subunit Vps32*	0	1	0	0	0	0	0	0	0	1	0
SPCC23B6.04c		*sec14 cytosolic factor family, glycerophospholipid-transfer protein (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPCC1322.14c	***vtc4***	*vacuolar transporter chaperone (VTC) complex subunit (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPCC1620.04c	***mug55***	*Cdc20/Fizzy subfamily WD repeat protein*	0	1	0	0	0	0	0	0	0	1	0
SPAP14E8.02	***tos4***	*FHA domain protein Tos4 (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPAC1A6.09c	***lag1***	*sphingosine N-acyltransferase Lag1*	0	1	0	0	0	0	0	0	0	1	0
SPCC1322.06	***kap113***	*karyopherin Kap113*	0	1	0	0	0	0	0	0	0	1	0
SPAC631.02	***bdf2***	*BET family double bromodomain protein Bdf2*	0	1	0	0	0	0	0	0	0	1	0
SPCC548.07c	***ght1***	*hexose transmembrane transporter Ght1*	0	1	0	0	0	0	0	0	0	1	0
SPAC343.10	***met11***	*methylenetetrahydrofolate reductase Met11*	0	1	0	0	0	0	0	0	0	1	0
SPCC1739.15	***wtf21***	*wtf element Wtf21*	0	1	0	0	0	0	0	0	0	1	0
SPAC25G10.06	***rps2801***	*40S ribosomal protein S28 (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPCC320.05		*sulfate transmembrane transporter (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPAC19G12.16c	***adg2***	*conserved fungal protein Adg2*	0	1	0	0	0	0	0	0	0	1	0
SPAC25B8.19c	***loz1***	*transcription factor zf-C2H2 type*	0	1	0	0	0	0	0	0	0	1	0
SPCC1739.09c	***cox13***	*cytochrome c oxidase subunit VIa (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPCC162.04c	***wtf13***	*wtf element Wtf13*	0	1	0	0	0	0	0	0	0	1	0
SPCC584.13		*amino acid permease (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPBC577.02	***rpl3801***	*60S ribosomal protein L38 (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPCC594.07c	***bqt3***	*bouquet formation protein Bqt3*	0	1	0	0	0	0	0	0	0	1	0
SPAC630.13c	***tsc2***	*tuberin*	0	1	0	0	0	0	0	0	0	1	0
SPCC188.13c	***dcr1***	*dicer*	0	1	0	0	0	0	0	0	0	1	0
SPAC144.06	***apl5***	*AP-3 adaptor complex subunit Apl5 (predicted)*	0	1	0	0	0	0	0	0	0	1	0
SPAC1F7.14c	***tam6***	*mitochondrial conserved protein*	0	1	0	0	0	0	0	0	0	0	1
SPBC887.10	***mcs4***	*response regulator Mcs4*	0	0	1	0	0	0	0	0	0	1	0	0
SPAC22E12.14c	***sck2***	*serine/threonine protein kinase Sck2*	0	0	1	0	0	0	0	0	0	1	0	0
SPCC663.12	***cid12***	*poly(A) polymerase Cid12*	0	0	1	0	0	0	0	0	0	1	0	0
SPAC2G11.09		*calcium ion transmembrane transporter (predicted)*	0	0	1	0	0	0	0	0	0	1	0	0
SPBC2D10.06	***rep1***	*MBF transcription factor activator Rep1*	0	0	1	0	0	0	0	0	0	0	1	0
SPBC8D2.10c	***rmt3***	*type I ribosomal protein arginine N-methyltransferase Rmt3*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC16E8.09	***scd1***	*RhoGEF Scd1*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC3H5.07	***rpl702***	*60S ribosomal protein L7*	0	0	1	0	0	0	0	0	0	0	1	0
SPCC338.07c	***naa15***	*NatA N-acetyltransferase complex regulatory subunit Naa15 (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPAP7G5.06	***per1***	*plasma membrane amino acid permease Per1*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC29A4.12c	***mug108***	*Schizosaccharomyces specific protein Mug108*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC6B12.07c		*ubiquitin-protein ligase E3 (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC3G9.05	***spa2***	*cell polarity protein Spa2*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC29A4.20	***elp3***	*elongator complex subunit Elp3 (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC30D11.12	***rpl3802***	*60S ribosomal protein L38 (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC630.14c	***tup12***	*transcriptional corepressor Tup12*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC25B8.10		*trans-aconitate 3-methyltransferase (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPCC1742.01	***gsf2***	*galactose-specific flocculin Gsf2*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC5H10.05c		*FAD binding oxidoreductase (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPCC1494.01		*iron/ascorbate oxidoreductase family*	0	0	1	0	0	0	0	0	0	0	1	0

SPAC3A11.07	***nde2***	*mitochondrial NADH dehydrogenase (ubiquinone) Nde2 (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC4H3.01		*DNAJ domain protein Caj1/Djp1 type (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPCC18.20		*dubious*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC4A8.05c	***myp2***	*myosin II heavy chain Myo3*	0	0	1	0	0	0	0	0	0	0	1	0
SPBC12D12.02c	***cdm1***	*DNA polymerase delta subunit Cdm1*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC19G12.04	***dal1***	*ureidoglycolate hydrolase (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPCC1259.04	***iec3***	*Ino80 complex subunit Iec3*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC8E11.03c	***dmc1***	*RecA family ATPase Dmc1*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC56F8.16	***esc1***	*transcription factor Esc1 (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC6C3.04	***cit1***	*citrate synthase Cit1*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC977.01		*S. pombe specific 5Tm protein family*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC29B12.02c	***set2***	*histone lysine methyltransferase Set2*	0	0	1	0	0	0	0	0	0	0	1	0
SPBC1778.01c	***zuo1***	*zuotin (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPBC2G2.13c	***dcd1***	*deoxycytidylate deaminase (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC1952.05	***gcn5***	*SAGA complex histone acetyltransferase catalytic subunit Gcn5*	0	0	1	0	0	0	0	0	0	0	1	0
SPCC1259.07	***rxt3***	*transcriptional regulatory protein Rxt3*	0	0	1	0	0	0	0	0	0	0	1	0
SPAC3G6.01	***hrp3***	*ATP-dependent DNA helicase Hrp3*	0	0	1	0	0	0	0	0	0	0	1	0
SPBC2G2.07c	***mug178***	*mitochondrial ribosomal protein subunit L51-b (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPBC947.15c	***nde1***	*mitochondrial NADH dehydrogenase (ubiquinone) Nde1 (predicted)*	0	0	1	0	0	0	0	0	0	0	1	0
SPCC24B10.08c	***ada2***	*SAGA complex subunit Ada2*	0	0	1	0	0	0	0	0	0	0	0	1
SPBC56F2.05c		*transcription factor (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC1259.02c	***erm1***	*Endoplasmic Reticulum metallopeptidase Erm1 (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC736.08	***cbf11***	*CBF1/Su(H)/LAG-1 family transcription factor Cbf11*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC777.08c	***bit61***	*TORC2 subunit Bit61*	0	0	1	0	0	0	0	0	0	0	0	1
SPAC1486.11	***fmc1***	*mitochondrial matrix protein, F1F0 ATP synthase assembly factor Fmc1 (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC970.06		*cargo receptor for soluble proteins (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPBC428.04	***apq12***	*nuclear membrane organization protein Apq12 (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC74.03c	***ssp2***	*AMP-activated protein serine/threonine kinase alpha subunit Ssp2*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC1450.12		*PXA domain protein*	0	0	1	0	0	0	0	0	0	0	0	1
SPCPJ732.02c	***xks1***	*xylulose kinase Xks1 (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPBC1921.03c	***mex67***	*mRNA export receptor, Tap, nucleoporin Mex67*	0	0	1	0	0	0	0	0	0	0	0	1
SPAC24H6.04	***hxk1***	*hexokinase 1*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC4G3.10c	***rhp42***	*DNA repair protein Rhp42*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC63.08c	***atg1***	*autophagy and CVT pathway serine/threonine protein kinase Atg1*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC794.09c	***tef101***	*translation elongation factor EF-1 alpha Ef1a-a*	0	0	1	0	0	0	0	0	0	0	0	1
SPAC30.04c	***abc4***	*glutathione S-conjugate-exporting ATPase Abc4*	0	0	1	0	0	0	0	0	0	0	0	1
SPAC4D7.14	***new13***	*conserved fungal protein of unknown function*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC18.13		*tRNA (guanine-N7-)-methyltransferase subunit Trm82 (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC70.03c		*proline dehydrogenase (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPBC3H7.08c		*conserved fungal protein*	0	0	1	0	0	0	0	0	0	0	0	1
SPCC777.12c		*thioredoxin family protein*	0	0	1	0	0	0	0	0	0	0	0	1
SPAC630.09c	***mug58***	*GLYK family kinase of unknown specificity (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPAC27D7.05c	***apc14***	*anaphase-promoting complex subunit Apc14*	0	0	1	0	0	0	0	0	0	0	0	1
SPAC22G7.08	***ppk8***	*serine/threonine protein kinase Ppk8 (predicted)*	0	0	1	0	0	0	0	0	0	0	0	1
SPCP1E11.03	***mug170***	*arrestin family Schizosaccharomyces specific protein Mug170*	0	0	1	0	0	0	0	0	0	0	0	1
SPBC409.03	***swi5***	*Swi5 protein*	0	0	1	0	0	0	0	0	0	0	0	1

SPBC215.03c	**csn1**	COP9/signalosome complex subunit Csn1	0	0	1	0	0	0	0	0	0	0	1
SPAC7D4.02c	**sfp47**	Ubp4 interactor Sfp47	0	0	1	0	0	0	0	0	0	0	1
SPAC23G3.10c	**ssr3**	SWI/SNF and RSC complex subunit Ssr3	0	0	1	0	0	0	0	0	0	0	1
SPAC22A12.11	**dak1**	dihydroxyacetone kinase Dak1	0	0	1	0	0	0	0	0	0	0	1
SPAC1A6.06c	**meu31**	Schizosaccharomyces specific protein Meu31	0	0	1	0	0	0	0	0	0	0	1
SPCC1322.09		conserved fungal protein	0	0	1	0	0	0	0	0	0	0	1
SPAC6F6.13c		DUF726 family protein	0	0	1	0	0	0	0	0	0	0	1
SPAC4D7.07c	**csi2**	mitotic chromosome segregation protein Csi2	0	0	1	0	0	0	0	0	0	0	1
SPAPB1A10.14	**pof15**	F-box protein (predicted)	0	0	1	0	0	0	0	0	0	0	1
SPBC13G1.12	**did2**	ESCRT III complex subunit Did2 (predicted)	0	0	1	0	0	0	0	0	0	0	1
SPAC19D5.06c	**din1**	Dhp1p-interacting protein Din1	0	0	1	0	0	0	0	0	0	0	1
SPAC607.07c		Schizosaccharomyces specific protein	0	0	1	0	0	0	0	0	0	0	1
SPCC1259.08		conserved fungal protein, DUF2457 family	0	0	1	0	0	0	0	0	0	0	1
SPAC4F10.08	**mug126**	Schizosaccharomyces pombe specific protein	0	0	1	0	0	0	0	0	0	0	1
SPBC18H10.15	**cdk11**	serine/threonine protein kinase cdk11	0	0	0	0	0	0	0	0	1	1	0
SPBC28E12.02		RNA-binding protein	0	0	0	0	0	0	0	0	1	1	0
SPBC18H10.10c	**saf4**	splicing associated factor Saf4	0	0	0	0	0	0	0	0	1	1	0
SPBC18H10.20c	**any1**	arrestin-related endocytic adaptor Any1	0	0	0	0	0	0	0	0	1	1	0
SPAC1071.03c	**sil1**	nucleotide exchange factor for the ER lumenal Hsp70 chaperone, Sil1 (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPBC3H7.06c	**pof9**	F-box protein Pof9	0	0	0	0	0	0	0	0	1	1	0
SPBC32F12.01c	**css1**	inositol phosphosphingolipid phospholipase C, Css1	0	0	0	0	0	0	0	0	1	1	0
SPBC21.07c	**ppk24**	serine/threonine protein kinase Ppk24	0	0	0	0	0	0	0	0	1	1	0
SPBC9B6.09c	**mdl1**	mitochondrial peptide-transporting ATPase	0	0	0	0	0	0	0	0	1	1	0
SPBC3H7.13	**far10**	SIP/FAR complex FHA domain subunit Far10/Csc1	0	0	0	0	0	0	0	0	1	1	0
SPBC11G11.02c	**end3**	actin cortical patch component End3 (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPCC16C4.20c	**hap2**	HMG box protein (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPCC4B3.04c	**nte1**	lysophospholipase (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPAC140.04	**ctr1**	conserved eukaryotic protein, human CCDC174 ortholog	0	0	0	0	0	0	0	0	1	1	0
SPBC32H8.13c	**mok12**	alpha-1,3-glucan synthase Mok12	0	0	0	0	0	0	0	0	1	1	0
SPBC18H10.19	**vps38**	phophatidylinositol 3-kinase complex subunit Vps38	0	0	0	0	0	0	0	0	1	1	0
SPAC18B11.09c		serine O-acetyltransferase activity (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPBC18H10.18c		Schizosaccharomyces specific protein	0	0	0	0	0	0	0	0	1	1	0
SPCC1393.09c	**gir2**	RWD domain protein, involved in cytoplasmic translation Gir2	0	0	0	0	0	0	0	0	1	1	0
SPBC18H10.05		WD repeat protein, human WDR44 family	0	0	0	0	0	0	0	0	1	1	0
SPBC28E12.06c	**lvs1**	beige protein homolog Lvs1 (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPBC27B12.09c		mitochondrial FAD transmembrane transporter (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPAC683.02c		zf-CCHC type zinc finger protein (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPAC1399.04c		uracil phosphoribosyltransferase (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPAC18G6.02c	**chp1**	chromodomain protein Chp1	0	0	0	0	0	0	0	0	1	1	0
SPAC12B10.12c	**rhp41**	DNA repair protein Rhp41	0	0	0	0	0	0	0	0	1	1	0
SPAC31G5.04	**lys12**	homoisocitrate dehydrogenase Lys12	0	0	0	0	0	0	0	0	1	1	0
SPBC1709.11c	**png2**	ING family homolog Png2	0	0	0	0	0	0	0	0	1	1	0
SPAC6F6.01	**cch1**	calcium ion channel Cch1	0	0	0	0	0	0	0	0	1	1	0
SPBC28E12.03	**rga4**	Rho-type GTPase activating protein Rga4	0	0	0	0	0	0	0	0	1	1	0
SPAC8F11.08c		esterase/lipase (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPAC31A2.14	**bun107**	WD repeat protein, human WDR48 family Bun107	0	0	0	0	0	0	0	0	1	1	0
SPBC365.13c	**hba1**	Ran GTPase binding protein Hba1	0	0	0	0	0	0	0	0	1	1	0
SPAC30.02c		elongator complex associated protein Kti2 (predicted)	0	0	0	0	0	0	0	0	1	1	0
SPAC23H4.12	**alp13**	MRG family Clr6 histone deacetylase complex subunit Alp13	0	0	0	0	0	0	0	0	1	1	0
SPBC3H7.14	**mug176**	BRCT domain protein	0	0	0	0	0	0	0	0	1	1	0

SPAC17A5.16	***ftp105***	*Ubp5 interacting protein Ftp105*	0	0	0	0	0	0	0	0	0	1	1	0
SPCC417.02	***dad5***	*DASH complex subunit Dad5*	0	0	0	0	0	0	0	0	0	1	1	0
SPBC1685.13	***fhn1***	*plasma membrane organization protein Fhn1*	0	0	0	0	0	0	0	0	0	1	1	0
SPBC1685.02c	***rps1202***	*40S ribosomal protein S12 (predicted)*	0	0	0	0	0	0	0	0	0	1	1	0
SPBC3H7.02		*sulfate transmembrane transporter (predicted)*	0	0	0	0	0	0	0	0	0	1	1	0
SPAC12B10.11	***exg2***	*glucan glucosidase Exg2, unknown specificity*	0	0	0	0	0	0	0	0	0	1	1	0
SPCC188.08c	***ubp5***	*ubiquitin C-terminal hydrolase Ubp5*	0	0	0	0	0	0	0	0	0	1	1	0
SPAC4F10.19c	***hit1***	*zf-HIT family C/D snoRNP assembly protein Hit1 (predicted)*	0	0	0	0	0	0	0	0	0	1	1	0
SPAC13A11.06		*pyruvate decarboxylase (predicted)*	0	0	0	0	0	0	0	0	0	1	1	0
SPBC11G11.01	***fis1***	*mitochondrial fission protein Fis1 (predicted)*	0	0	0	0	0	0	0	0	0	1	1	0
SPAC8E11.05c		*conserved fungal protein, associated with clathrin coated vesicles (predicted)*	0	0	0	0	0	0	0	0	0	1	1	0
SPAC22H12.04c	***rps102***	*40S ribosomal protein S3a (predicted)*	0	0	0	0	0	0	0	0	0	1	1	0
SPAC13C5.04		*amidotransferase (predicted)*	0	0	0	0	0	0	0	0	0	1	1	0
SPAC824.04	***swd22***	*mRNA cleavage and polyadenylation specificity factor complex subunit, WD repeat protein Swd22*	0	0	0	0	0	0	0	0	0	1	1	0
SPBC20F10.07		*GRAM domain protein*	0	0	0	0	0	0	0	0	0	1	1	0
SPAC17C9.05c	***pmc3***	*mediator complex subunit Med27*	0	0	0	0	0	0	0	0	0	1	1	0
SPBC336.01	***fbh1***	*DNA helicase I, ubiquitin ligase F-box adaptor Fbh1*	0	0	0	0	0	0	0	0	0	1	0	1
SPBC16H5.13		*WD repeat protein, human WDR7 ortholog*	0	0	0	0	0	0	0	0	0	1	0	1
SPAC823.13c		*mitochondrial inner membrane protein (predicted)*	0	0	0	0	0	0	0	0	0	1	0	1
SPAC19A8.01c	***sec73***	*guanyl-nucleotide exchange factor Sec73 (predicted)*	0	0	0	0	0	0	0	0	0	0	1	1
SPBC4B4.08	***ght2***	*hexose transmembrane transporter Ght2*	0	0	0	0	0	0	0	0	0	0	1	1
SPCC320.07c	***mde7***	*RNA-binding protein Mde7*	0	0	0	0	0	0	0	0	0	0	1	1
SPBC16E9.11c	***pub3***	*HECT-type ubiquitin-protein ligase E3 Pub3 (predicted)*	0	0	0	0	0	0	0	0	0	0	1	1
SPCC1259.09c	***pdx1***	*pyruvate dehydrogenase protein x component, Pdx1 (predicted)*	0	0	0	0	0	0	0	0	0	0	1	1
SPAC977.15		*dienelactone hydrolase family*	0	0	0	0	0	0	0	0	0	0	1	1
SPAC29A4.17c		*mitochondrial FUN14 family protein involved in mitophagy*	0	0	0	0	0	0	0	0	0	0	1	1
SPCC1682.12c	***ubp16***	*ubiquitin C-terminal hydrolase Ubp16*	0	0	0	0	0	0	0	0	0	0	1	1
SPBP35G2.05c	***cki2***	*serine/threonine protein kinase Cki2*	0	0	0	0	0	0	0	0	0	0	1	1
SPBP23A10.16	***sdh4***	*TIM22 inner membrane protein import complex anchor subunit Tim18*	0	0	0	0	0	0	0	0	0	0	1	1

Appendix D: Variants identified in the strains used

Chr	Position	Reference sequence	Alternate sequence	Type	Pem2 Replicate 1	Pem2 Replicate 2	ell1Δ Replicate 1	ell1Δ Replicate 2	eaf1Δ Replicate 1	eaf1Δ Replicate 2	ebp1Δ Replicate 1	ebp1Δ Replicate 2	Remarks
I	440084	T	TA	INDEL	TA	T	TA	T	TA	T	TA	TA	*TTTTT**TA**AAAAAAAAAAAAA
I	453369	G	GAA	INDEL	GAA	.	.	GAA	.	GAA	.	GAA	***GAA**AAAAAAAAAAAAA
I	650099	C	CAA	INDEL	CAA	C	CAA	CAA	.	C	CAA	.	***CAA**AAAAAAAAAAAAA
I	660056	T	TG	INDEL	TG	TG	TG	T	TG	TG	TG	TG	***TG**GGGGGGGGG
I	960025	G	GT	INDEL	GT	GT	GT	GT	GT	GT	GT	GT	***GT**TTTTTTTTTT
I	965438	C	CA	INDEL	CA	CA	CA	CA	.	C	.	CA	***CA**AAAAAAAAAAAAA
I	1358794	C	CAA	INDEL	CAA	CAA	CAA	CAA	CAA	CAA	CAA	CAA	*C**CA**AAAAAAAAAAAA
I	1595056	C	CT	INDEL	CT	CT	C	C	.	CT	CT	.	***CT**TTTTTTTTTTTTTT
I	1905525	T	TG	INDEL	TG	TG	TG	.	T	TG	T	TG	*TTTTTTTT**TG**GGGGGGGG
I	1955853	T	TA	INDEL	TA	TA	TA	TA	TA	TA	TA	TA	***TA**AAAAAAAAAAAAA
I	3166776	G	GTTT	INDEL	G	GTTT	GTTT	GTTT	GTTT	GTTT	GTTT	GTTT	***GT**TTTTTTTTTTTTT
I	3659724	C	CT	INDEL	CT	C	C	C	C	CT	CT	CT	***CT**TTTTTTTTTTTTT
I	4916623	C	CA	INDEL	.	C	C	.	C	CA	CA	C	***CA**AAAAAAAAAAAAA
I	5176926	T	TA	INDEL	TA	TA	T	T	TA	TA	TA	T	***TA**AAAAAAAAAAAAA
I	5426622	C	CT	INDEL	CT	CT	CT	CT	CT	CT	C	CT	***CT**TTTTTTTTTT
II	889921	C	CTTTT	INDEL	C	C	CTTTT	CTTTT	C	C	C	CTTTT	***CT**TTTTTTTTTTTTT
II	1266289	T	TA	INDEL	.	TA	TA	TA	.	TA	.	TA	***TA**AAAAAAAAAAAAA
II	1555919	A	ATTTATT	INDEL	ATTTATT	ATTTATT	ATTTATT	ATTTATT	ATTTATT	ATTTATT	.	ATTTATT	***ATTTATT**TTTATTTTT
II	1716256	G	GT	INDEL	G	G	G	GT	GT	GT	G	GT	***GT**TTTTTTTTTTT
II	1867795	A	AT	INDEL	AT	AT	.	AT	AT	AT	AT	AT	***AT**TTTTTTTTT
II	1868122	G	GTT	INDEL	GTT	GTT	GTT	GTT	G	GTT	GTT	G	***GTT**TTTTTTTTTTTTT
II	2630607	C	CA	INDEL	CA	CA	C	CA	CA	CA	CA	CA	***CA**AAAAAAAAAAAAA
II	2640785	T	TA	INDEL	.	TA	TA	TA	T	T	TA	T	*T**TA**AAAAAAAAAAAAA
II	2755338	CAT	C	INDEL	C	C	CAT	C	C	C	C	C	***CAT**ATATATATATATATAT
II	3799350	G	GA	INDEL	GA	GA	G	GA	GA	G	GA	GA	***GA**AAAAAAAAAAAAA
II	4233001	C	CTA	INDEL	C	C	C	C	C	C	CTA	CTA	***CTA**TATATATATATA
II	4419921	T	TA	INDEL	.	TA	.	TA	.	T	.	TA	*T**TA**AAAAAAAAAAAAA
III	107118	G	C	SNP	G	G	G	G	G	G	G	C	SNP in only 1 of the 3 replicates
III	719971	G	T	SNP	T	T	T	T	T	T	.	T	variant in all strains used
III	1168363	G	A	SNP	A	A	G	A	A	G	G	A	*GG**GA**AAAAAAAAAAAA
III	1861945	A	AC	INDEL	A	A	A	A	AC	AC	A	A	Eaf1 gene position
III	1861966	TA	T	INDEL	TA	TA	TA	TA	T	T	TA	TA	Eaf1 gene position

The Wildtype and mutant strains (all in Pem2 background) were sequenced and aligned to the reference *S. pombe* sequence, in replicates. This table lists all the variants (INDELs and SNPs) that were identified. The alternate sequence refers to the sequence obtained that were different from the refence sequence. Remarks section for most of the rows show the sequence at the region with the base-pairs at the position in bold. (* denotes that many of the apparent variants are in regions with long runs of the same residue where sequencing errors may be more common.)

References

Ahn,S.H., Kim,M., and Buratowski,S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Mol. Cell *13*, 67-76.

Akhtar,M.S., Heidemann,M., Tietjen,J.R., Zhang,D.W., Chapman,R.D., Eick,D., and Ansari,A.Z. (2009). TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. Mol. Cell *34*, 387-393.

Allshire,R.C. and Ekwall,K. (2015). Epigenetic Regulation of Chromatin States in Schizosaccharomyces pombe. Cold Spring Harb. Perspect. Biol. *7*, a018770.

Aravind,L., Watanabe,H., Lipman,D.J., and Koonin,E.V. (2000). Lineage-specific loss and divergence of functionally linked genes in eukaryotes. Proc. Natl. Acad. Sci. U. S. A *97*, 11319-11324.

Asturias,F.J., Jiang,Y.W., Myers,L.C., Gustafsson,C.M., and Kornberg,R.D. (1999). Conserved structures of Mediator and RNA polymerase II holoenzyme. Science *283*, 985-987.

Bahler,J., Wu,J.Q., Longtine,M.S., Shah,N.G., McKenzie III,A., Steever,A.B., Wach,A., Philippsen,P., and Pringle,J.R. (1998). Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast *14*, 943-951.

Banks,C.A., Kong,S.E., Spahr,H., Florens,L., Martin-Brown,S., Washburn,M.P., Conaway,J.W., Mushegian,A., and Conaway,R.C. (2007). Identification and Characterization of a Schizosaccharomyces pombe RNA Polymerase II Elongation Factor with Similarity to the Metazoan Transcription Factor ELL. J. Biol. Chem. *282*, 5761-5769.

Baryshnikova,A., Costanzo,M., Dixon,S., Vizeacoumar,F.J., Myers,C.L., Andrews,B., and Boone,C. (2010). Synthetic genetic array (SGA) analysis in Saccharomyces cerevisiae and Schizosaccharomyces pombe. Methods Enzymol. *470*, 145-179.

Baryshnikova,A., Costanzo,M., Myers,C.L., Andrews,B., and Boone,C. (2013). Genetic interaction networks: toward an understanding of heritability. Annu. Rev. Genomics Hum. Genet. *14*, 111-133.

Beve,J., Hu,G.Z., Myers,L.C., Balciunas,D., Werngren,O., Hultenby,K., Wibom,R., Ronne,H., and Gustafsson,C.M. (2005a). The structural and functional role of Med5 in the yeast Mediator tail module. J. Biol. Chem. *280*, 41366-41372.

Beve,J., Hu,G.Z., Myers,L.C., Balciunas,D., Werngren,O., Hultenby,K., Wibom,R., Ronne,H., and Gustafsson,C.M. (2005b). The structural and functional role of Med5 in the yeast Mediator tail module. J. Biol. Chem. *280*, 41366-41372.

Biswas,D., Milne,T.A., Basrur,V., Kim,J., Elenitoba-Johnson,K.S., Allis,C.D., and Roeder,R.G. (2011). Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. Proc. Natl. Acad. Sci. U. S. A *108*, 15751-15756.

Bitoun,E., Oliver,P.L., and Davies,K.E. (2007). The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. Hum. Mol. Genet. *16*, 92-106.

Boeing,S., Rigault,C., Heidemann,M., Eick,D., and Meisterernst,M. (2010). RNA polymerase II C-terminal heptarepeat domain Ser-7 phosphorylation is established in a mediator-dependent fashion. J. Biol. Chem. *285*, 188-196.

Booth,G.T., Wang,I.X., Cheung,V.G., and Lis,J.T. (2016). Divergence of a conserved elongation factor and transcription regulation in budding and fission yeast. Genome Res. *26*, 799-811.

Bourbon,H.M. (2008). Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. Nucleic Acids Res. *36*, 3993-4008.

Bradsher,J.N., Tan,S., McLaury,H.-J., Conaway,J.W., and Conaway,R.C. (1993). RNA polymerase II transcription factor SIII: II. Functional properties and role in RNA chain elongation. J. Biol. Chem. *268*, 25594-25603.

Buratowski,S. (2009). Progression through the RNA polymerase II CTD cycle. Mol. Cell *36*, 541-546.

Byun,J.S., Fufa,T.D., Wakano,C., Fernandez,A., Haggerty,C.M., Sung,M.H., and Gardner,K. (2012). ELL facilitates RNA polymerase II pause site entry and release. Nat. Commun. *3*, 633.

Chen,H.M., Rosebrock,A.P., Khan,S.R., Futcher,B., and Leatherwood,J.K. (2012). Repression of meiotic genes by antisense transcription and by Fkh2 transcription factor in Schizosaccharomyces pombe. PLoS. One. *7*, e29917.

Chen,Y., Zhou,C., Ji,W., Mei,Z., Hu,B., Zhang,W., Zhang,D., Wang,J., Liu,X., Ouyang,G., Zhou,J., and Xiao,W. (2016). ELL targets c-Myc for proteasomal degradation and suppresses tumour growth. Nat. Commun. *7*, 11057.

Chou,S., Upton,H., Bao,K., Schulze-Gahmen,U., Samelson,A.J., He,N., Nowak,A., Lu,H., Krogan,N.J., Zhou,Q., and Alber,T. (2013). HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold. Proc. Natl. Acad. Sci. U. S. A *110*, E123-E131.

Clarke,L. (1990). Centromeres of budding and fission yeasts. Trends Genet. *6*, 150-154.

Conaway,J.W. and Conaway,R.C. (1997). General transcription factors for RNA polymerase II. Prog. Nucleic. Acids. Res. Mol. Biol. *56*, 327-346.

Conaway,R.C. and Conaway,J.W. (2013). The Mediator complex and transcription elongation. Biochim. Biophys. Acta *1829*, 69-75.

Coudreuse,D., van,B.H., Dewez,M., Soutourina,J., Parnell,T., Vandenhaute,J., Cairns,B., Werner,M., and Hermand,D. (2010a). A gene-specific requirement of RNA polymerase II CTD phosphorylation for sexual differentiation in S. pombe. Curr. Biol. *20*, 1053-1064.

Coudreuse,D., van,B.H., Dewez,M., Soutourina,J., Parnell,T., Vandenhaute,J., Cairns,B., Werner,M., and Hermand,D. (2010b). A gene-specific requirement of RNA polymerase II CTD phosphorylation for sexual differentiation in S. pombe. Curr. Biol. *20*, 1053-1064.

Dabas,P., Sweta,K., Ekka,M., and Sharma,N. (2018). Structure function characterization of the ELL Associated Factor (EAF) from Schizosaccharomyces pombe. Gene *641*, 117-128.

Dheur,S., Saupe,S.J., Genier,S., Vazquez,S., and Javerzat,J.P. (2011). Role for cohesin in the formation of a heterochromatic domain at fission yeast subtelomeres. Mol. Cell Biol. *31*, 1088-1097.

DiMartino,J.F., Miller,T., Ayton,P.M., Landewe,T., Hess,J.L., Cleary,M.L., and Shilatifard,A. (2000). A carboxy-terminal domain of ELL is required and sufficient for immortalization of myeloid precursors by MLL-ELL. Blood *96*, 3887-3893.

Donner,A.J., Ebmeier,C.C., Taatjes,D.J., and Espinosa,J.M. (2010). CDK8 is a positive regulator of transcriptional elongation within the serum response network. Nat. Struct. Mol. Biol. *17*, 194-201.

Donner,A.J., Szostek,S., Hoover,J.M., and Espinosa,J.M. (2007). CDK8 is a stimulus-specific positive coregulator of p53 target genes. Mol. Cell *27*, 121-133.

Dotson,M.R., Yuan,C.X., Roeder,R.G., Myers,L.C., Gustafsson,C.M., Jiang,Y.W., Li,Y., Kornberg,R.D., and Asturias,F.J. (2000). Structural organization of yeast and mammalian mediator complexes. Proc. Natl. Acad. Sci. U. S. A. *97*, 14307-14310.

Dover,J., Schneider,J., Tawiah-Boateng,M.A., Wood,A., Dean,K., Johnston,M., and Shilatifard,A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J. Biol. Chem. *277*, 28368-28371.

Drogat,J., Migeot,V., Mommaerts,E., Mullier,C., Dieu,M., van,B.H., and Hermand,D. (2012). Cdk11-cyclinL controls the assembly of the RNA polymerase II mediator complex. Cell Rep. *2*, 1068-1076.

Dvir,A., Conaway,J.W., and Conaway,R.C. (2001). Mechanism of transcription initiation and promoter escape by RNA polymerase II. Curr. Opin. Genet. Dev. *11*, 209-214.

Eissenberg,J.C., Ma,J., Gerber,M.A., Christensen,A., Kennison,J.A., and Shilatifard,A. (2002). dELL is an essential RNA polymerase II elongation factor with a general role in development. Proc. Natl. Acad. Sci. U. S. A. *99*, 9894-9899.

Eissenberg,J.C., Shilatifard,A., Dorokhov,N., and Michener,D.E. (2007). Cdk9 is an essential kinase in Drosophila that is required for heat shock gene expression, histone methylation and elongation factor recruitment. Molecular Genetics and Genomics *277*, 101-114.

Ekwall,K., Nimmo,E.R., Javerzat,J.P., Borgstrom,B., Egel,R., Cranston,G., and Allshire,R. (1996). Mutations in the fission yeast silencing factors clr4+ and rik1+ disrupt the localisation of the chromo domain protein Swi6p and impair centromere function. J. Cell Sci. *109 (Pt 11)*, 2637-2648.

Ekwall,K. and Ruusala,T. (1994). Mutations in rik1, clr2, clr3 and clr4 genes asymmetrically derepress the silent mating-type loci in fission yeast. Genetics *136*, 53-64.

Elmendorf,B.J., Shilatifard,A., Yan,Q., Conaway,J.W., and Conaway,R.C. (2001). Transcription Factors TFIIF, ELL, and Elongin Negatively Regulate SII- induced Nascent Transcript Cleavage by Non-arrested RNA Polymerase II Elongation Intermediates. J. Biol. Chem. *276*, 23109-23114.

Elmlund,H., Baraznenok,V., Lindahl,M., Samuelsen,C.O., Koeck,P.J., Holmberg,S., Hebert,H., and Gustafsson,C.M. (2006). The cyclin-dependent kinase 8 module sterically blocks Mediator interactions with RNA polymerase II. Proc. Natl. Acad. Sci. U. S. A *103*, 15788-15793.

Florens,L., Carozza,M.J., Swanson,S.K., Fournier,M., Coleman,M.K., Workman,J.L., and Washburn,M.P. (2006). Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. Methods *40*, 303-311.

Florens,L. and Washburn,M.P. (2006). Proteomic analysis by multidimensional protein identification technology. Methods Mol. Biol. *328*, 159-175.

Fuchs,G., Hollander,D., Voichek,Y., Ast,G., and Oren,M. (2014). Cotranscriptional histone H2B monoubiquitylation is tightly coupled with RNA polymerase II elongation rate. Genome Res. *24*, 1572-1583.

Furumoto,T., Tanaka,A., Ito,M., Malik,S., Hirose,Y., Hanaoka,F., and Ohkuma,Y. (2007). A kinase subunit of the human mediator complex, CDK8, positively regulates transcriptional activation. Genes Cells *12*, 119-132.

Gall,J.G., Bellini,M., Wu,Z., and Murphy,C. (1999). Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. Mol. Biol. Cell *10*, 4385-4402.

Gerber,M., Ma,J., Dean,K., Eissenberg,J.C., and Shilatifard,A. (2001). Drosophila ELL is associated with actively elongating RNA polymerase II on transcriptionally active sites in vivo. EMBO J. *20*, 6104-6114.

Gnatt,A.L., Cramer,P., Fu,J., Bushnell,D.A., and Kornberg,R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3A resolution. Science *292*, 1876-1882.

Green,S.R. and Johnson,A.D. (2004). Promoter-dependent roles for the Srb10 cyclin-dependent kinase and the Hda1 deacetylase in Tup1-mediated repression in Saccharomyces cerevisiae. Mol. Biol. Cell *15*, 4191-4202.

Harreman,M., Taschner,M., Sigurdsson,S., Anindya,R., Reid,J., Somesh,B., Kong,S.E., Banks,C.A., Conaway,R.C., Conaway,J.W., and Svejstrup,J.Q. (2009). Distinct ubiquitin ligases act sequentially for RNA polymerase II poly-ubiquitylation. Proc. Natl. Acad. Sci. U. S. A. *106*, 20705-20710.

He,N., Chan,C.K., Sobhian,B., Chou,S., Xue,Y., Liu,M., Alber,T., Benkirane,M., and Zhou,Q. (2011). Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. Proc. Natl. Acad. Sci. U. S. A *108*, E636-E645.

He,N., Liu,M., Hsu,J., Xue,Y., Chou,S., Burlingame,A., Krogan,N.J., Alber,T., and Zhou,Q. (2010). HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. Mol. Cell *38*, 428-438.

Hedges,S.B. (2002). The origin and evolution of model organisms. Nat. Rev. Genet. *3*, 838-849.

Ho,C.K. and Shuman,S. (1999). Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. Mol. Cell *3*, 405-411.

Hoffman,C.S., Wood,V., and Fantes,P.A. (2015). An Ancient Yeast for Young Geneticists: A Primer on the Schizosaccharomyces pombe Model System. Genetics *201*, 403-423.

Hong,E.J., Villen,J., Gerace,E.L., Gygi,S.P., and Moazed,D. (2005). A cullin E3 ubiquitin ligase complex associates with Rik1 and the Clr4 histone H3-K9 methyltransferase and is required for RNAi-mediated heterochromatin formation. RNA. Biol. *2*, 106-111.

Hsin,J.P., Sheth,A., and Manley,J.L. (2011). RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. Science *334*, 683-686.

Hu,D., Smith,E.R., Garruss,A.S., Mohaghegh,N., Varberg,J.M., Lin,C., Jackson,J., Gao,X., Saraf,A., Florens,L., Washburn,M.P., Eissenberg,J.C., and Shilatifard,A. (2013). The little elongation complex functions at initiation and elongation phases of snRNA gene transcription. Mol. Cell *51*, 493-505.

Inada,M., Nichols,R.J., Parsa,J.Y., Homer,C.M., Benn,R.A., Hoxie,R.S., Madhani,H.D., Shuman,S., Schwer,B., and Pleiss,J.A. (2016). Phospho-site mutants of the RNA Polymerase II C-terminal domain alter subtelomeric gene expression and chromatin modification state in fission yeast. Nucleic Acids Res. *44*, 9180-9189.

Ivanova,A.V., Bonaduce,M.J., Ivanov,S.V., and Klar,A.J. (1998). The chromo and SET domains of the Clr4 protein are essential for silencing in fission yeast. Nat. Genet. *19*, 192-195.

Jaehning,J.A. (2010). The Paf1 complex: platform or player in RNA polymerase II transcription? Biochim. Biophys. Acta *1799*, 379-388.

Jishage,M., Malik,S., Wagner,U., Uberheide,B., Ishihama,Y., Hu,X., Chait,B.T., Gnatt,A., Ren,B., and Roeder,R.G. (2012). Transcriptional regulation by Pol II(G) involving mediator and competitive interactions of Gdown1 and TFIIF with Pol II. Mol. Cell *45*, 51-63.

Kanamitsu,K. and Ikeda,S. (2011). Fission yeast homologs of human XPC and CSB, rhp41 and rhp26, are involved in transcription-coupled repair of methyl methanesulfonate-induced DNA damage. Genes Genet. Syst. *86*, 83-91.

Kaufer,N.F. and Potashkin,J. (2000). Analysis of the splicing machinery in fission yeast: a comparison with budding yeast and mammals. Nucleic Acids Res. *28*, 3003-3010.

Kelleher,R.J., Flanagan,P.M., and Kornberg,R.D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. Cell *61*, 1209-1215.

Kellinger,M.W., Ulrich,S., Chong,J., Kool,E.T., and Wang,D. (2012). Dissecting chemical interactions governing RNA polymerase II transcriptional fidelity. J. Am. Chem. Soc. *134*, 8231-8240.

Khorosjutina,O., Wanrooij,P.H., Walfridsson,J., Szilagyi,Z., Zhu,X., Baraznenok,V., Ekwall,K., and Gustafsson,C.M. (2010). A chromatin-remodeling protein is a component of fission yeast mediator. J. Biol. Chem. *285*, 29729-29737.

Kim,E., Du,L., Bregman,D.B., and Warren,S.L. (1997). Splicing factors associate with hyperphosphorylated RNA polymerase II in the absence of Pre-mRNA. J. Cell Biol. *136*, 19-28.

Kim,Y.J., Bjorklund,S., Li,Y., Sayre,M.H., and Kornberg,R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell *77*, 599-608.

Kitts,P.A. and Possee,R.D. (1993). A method for producing recombinant baculovirus expression vectors at high frequency. BioTechniques. *14*, 810-817.

Klug,A. (2001). A marvellous machine for making messages. Science *292*, 1844-1846.

Knuesel,M.T., Meyer,K.D., Bernecky,C., and Taatjes,D.J. (2009). The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. Genes Dev. *23*, 439-451.

Komarnitsky,P.B., Cho,E.-J., and Buratowski,S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. *14*, 2452-2460.

Kong,S.E., Banks,C.A.S., Shilatifard,A., Conaway,J.W., and Conaway,R. (2005). ELL-associated factors 1 and 2 are positive regulators of RNA polymerase II elongation factor ELL. Proceedings of the National Academy of Sciences of the United States of America *102*, 10094-10098.

Kostrewa,D., Zeller,M.E., Armache,K.J., Seizl,M., Leike,K., Thomm,M., and Cramer,P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature *462*, 323-330.

Kwak,H., Fuda,N.J., Core,L.J., and Lis,J.T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science *339*, 950-953.

Lavau,C., Luo,R.T., Du,C., and Thirman,M.J. (2000). Retrovirus-mediated gene transfer of MLL-ELL transforms primary myeloid progenitors and causes acute myeloid leukemias in mice. Proc. Natl. Acad. Sci. U. S. A *97*, 10984-10989.

Liao,S.M., Zhang,J., Jeffery,D.A., Koleske,A.J., Thompson,C.M., Chao,D.M., Viljoen,M., van Vuuren,H.J.J., and Young,R.A. (1995). A kinase-cyclin pair in the RNA polymerase II holoenzyme. Nature *374*, 193-196.

Lin,C., Smith,E.R., Takahashi,H., Lai,K.C., Martin-Brown,S., Florens,L., Washburn,M.P., Conaway,J.W., Conaway,R.C., and Shilatifard,A. (2010). AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Mol. Cell *37*, 429-437.

Linder,T., Rasmussen,N.N., Samuelsen,C.O., Chatzidaki,E., Baraznenok,V., Beve,J., Henriksen,P., Gustafsson,C.M., and Holmberg,S. (2008a). Two conserved modules of Schizosaccharomyces pombe Mediator regulate distinct cellular pathways. Nucleic Acids Res. *36*, 2489-2504.

Linder,T., Rasmussen,N.N., Samuelsen,C.O., Chatzidaki,E., Baraznenok,V., Beve,J., Henriksen,P., Gustafsson,C.M., and Holmberg,S. (2008b). Two conserved modules of Schizosaccharomyces pombe Mediator regulate distinct cellular pathways. Nucleic Acids Res. *36*, 2489-2504.

Lis,J.T., Mason,P., Peng,J., Price,D.H., and Werner,J. (2000). P-TEFb kinase recruitment and function at heat shock loci. Genes Dev. *14*, 792-803.

Lu,X., Zhu,X., Li,Y., Liu,M., Yu,B., Wang,Y., Rao,M., Yang,H., Zhou,K., Wang,Y., Chen,Y., Chen,M., Zhuang,S., Chen,L.F., Liu,R., and Chen,R. (2016). Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II. Nucleic Acids Res. *44*, 6853-6867.

Luo,R.T., Lavau,C., Du,C., Simone,F., Polak,P.E., Kawamata,S., and Thirman,M.J. (2001). The elongation domain of ELL is dispensable but its ELL-associated factor 1 interaction domain is essential for MLL-ELL-induced leukemogenesis. Mol. Cell. Biol. *21*, 5678-5687.

Luo,Z., Lin,C., and Shilatifard,A. (2012). The super elongation complex (SEC) family in transcriptional control. Nat. Rev. Mol. Cell Biol. *13*, 543-547.

Mahat,D.B., Kwak,H., Booth,G.T., Jonkers,I.H., Danko,C.G., Patel,R.K., Waters,C.T., Munson,K., Core,L.J., and Lis,J.T. (2016). Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). Nat. Protoc. *11*, 1455-1476.

Malik,S., Barrero,M.J., and Jones,T. (2007). Identification of a regulator of transcription elongation as an accessory factor for the human Mediator coactivator. Proc. Natl. Acad. Sci. U. S. A *104*, 6182-6187.

Mannini,L., Lamaze,C., Cucco,F., Amato,C., Quarantotti,V., Rizzo,I.M., Krantz,I.D., Bilodeau,S., and Musio,A. (2015). Mutant cohesin affects RNA polymerase II regulation in Cornelia de Lange syndrome. Sci. Rep. *5*, 16803.

Marshall,N.F., Peng,J., Xie,Z., and Price,D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J. Biol. Chem. *271*, 27176-27183.

Martincic,K., Alkan,S.A., Cheatle,A., Borghesi,L., and Milcarek,C. (2009). Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. Nat. Immunol. *10*, 1102-1109.

Matsuyama,A., Arai,R., Yashiroda,Y., Shirai,A., Kamata,A., Sekido,S., Kobayashi,Y., Hashimoto,A., Hamamoto,M., Hiraoka,Y., Horinouchi,S., and Yoshida,M. (2006). ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. Nat. Biotechnol. *24*, 841-847.

Maurus,D., Héligon,C., Bürger-Schwärzler,A., Brändlli,A.W., and Kühl,M. (2005). Noncanonical Wnt-4 signaling and EAF2 are required for eye development in *Xenopus laevis*. EMBO J. *24*, 1181-1191.

Meinhart,A. and Cramer,P. (2004). Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. Nature *430*, 223-226.

Milcarek,C., Albring,M., Langer,C., and Park,K.S. (2011). The eleven-nineteen lysine-rich leukemia gene (ELL2) influences the histone H3 protein modifications accompanying the shift to secretory immunoglobulin heavy chain mRNA production. J. Biol. Chem. *286*, 33795-33803.

Miller,T., Williams,K., Johnstone,R.W., and Shilatifard,A. (2000). Identification, cloning, expression, and biochemical characterization of the testis-specific RNA polymerase II elongation factor ELL3. J. Biol. Chem. *275*, 32052-32056.

Mohan,M., Lin,C., Guest,E., and Shilatifard,A. (2010). Licensed to elongate: a molecular mechanism for MLL-based leukaemogenesis. Nat. Rev. Cancer *10*, 721-728.

Moser,B.A., Raguimova,O.N., and Nakamura,T.M. (2015). Ccq1-Tpz1TPP1 interaction facilitates telomerase and SHREC association with telomeres in fission yeast. Mol. Biol. Cell *26*, 3857-3866.

Motamedi,M.R., Verdel,A., Colmenares,S.U., Gerber,S.A., Gygi,S.P., and Moazed,D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789-802.

Mourgues,S., Gautier,V., Lagarou,A., Bordier,C., Mourcet,A., Slingerland,J., Kaddoum,L., Coin,F., Vermeulen,W., Gonzales de,P.A., Monsarrat,B., Mari,P.O., and Giglia-Mari,G. (2013a). ELL, a novel TFIIH partner, is involved in transcription restart after DNA repair. Proc. Natl. Acad. Sci. U. S. A *110*, 17927-17932.

Mourgues,S., Gautier,V., Lagarou,A., Bordier,C., Mourcet,A., Slingerland,J., Kaddoum,L., Coin,F., Vermeulen,W., Gonzales de,P.A., Monsarrat,B., Mari,P.O., and Giglia-Mari,G. (2013b). ELL, a novel TFIIH partner, is involved in transcription restart after DNA repair. Proc. Natl. Acad. Sci. U. S. A *110*, 17927-17932.

Mueller,D., Bach,C., Zeisig,D., Garcia-Cuellar,M.P., Monroe,S., Sreekumar,A., Zhou,R., Nesvizhskii,A., Chinnaiyan,A., Hess,J.L., and Slany,R.K. (2007). A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. Blood *110*, 4445-4454.

Mueller,D., Garcia-Cuellar,M.P., Bach,C., Buhl,S., Maethner,E., and Slany,R.K. (2009). Misguided transcriptional elongation causes mixed lineage leukemia. PLoS. Biol. *7*, e1000249.

Myers,L.C., Gustafsson,C.M., Hayashibara,K.C., Brown,P.O., and Kornberg,R.D. (1999). Mediator protein mutations that selectively abolish activated transcription. Proc. Natl. Acad. Sci. U. S. A. *96*, 67-72.

Myers,L.C. and Kornberg,R.D. (2000). Mediator of Transcriptional Regulation. Annu. Rev. Biochem. *69*, 729-749.

Ni,T., Tu,K., Wang,Z., Song,S., Wu,H., Xie,B., Scott,K.C., Grewal,S.I., Gao,Y., and Zhu,J. (2010). The prevalence and regulation of antisense transcripts in Schizosaccharomyces pombe. PLoS. One. *5*, e15271.

Ni,Z., Schwarts,B.E., Werner,J., Suarez,J.-R., and Lis,T. (2004). Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes. Mol. Cell *13*, 55-65.

Nogi,Y. and Fukasawa,T. (1980). A novel mutation that affects utilization of galactose in Saccharomyces cerevisiae. Curr. Genet. *2*, 115-120.

Osley,M.A. (2004). H2B ubiquitylation: the end is in sight. Biochim. Biophys. Acta *1677*, 74-78.

Pal,M. and Luse,D.S. (2003). The initiation-elongation transition: Lateral mobility of RNA in RNA polymerase II complexes is greatly reduced at +8/+9 and absent by +23. PNAS *100*, 5700-5705.

Park,J.M., Werner,J., Kim,J.M., Lis,J.T., and Kim,Y.J. (2001). Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock. Mol. Cell *8*, 9-19.

Pei,Y., Schwer,B., and Shuman,S. (2003). Interactions between fission yeast Cdk9, Its cyclin partner Pch1, and mRNA capping enzyme Pct1 suggest an elongation checkpoint for mRNA quality control. J. Biol. Chem. *278*, 7180-7188.

Perales,R. and Bentley,D. (2009). "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. Mol. Cell *36*, 178-191.

Peterlin,B.M. and Price,D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. Mol. Cell *23*, 297-305.

Racine,A., Page,V., Nagy,S., Grabowski,D., and Tanny,J.C. (2012). Histone H2B ubiquitylation promotes activity of the intact Set1 histone methyltransferase complex in fission yeast. J. Biol. Chem. *287*, 19040-19047.

Rasmussen,E.B. and Lis,J.T. (1993). In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. Proc. Natl. Acad. Sci. U. S. A *90*, 7923-7927.

Reines,D. (2003). Use of RNA yeast polymerase II mutants in studying transcription elongation. Methods Enzymol. *371*, 284-292.

Ribar,B., Prakash,L., and Prakash,S. (2007). ELA1 and CUL3 are required along with ELC1 for RNA polymerase II polyubiquitylation and degradation in DNA-damaged yeast cells. Mol. Cell Biol. *27*, 3211-3216.

Richard,P. and Manley,J.L. (2009). Transcription termination by nuclear RNA polymerases. Genes Dev. *23*, 1247-1269.

Riles,L., Shaw,R.J., Johnston,M., and Reines,D. (2004). Large-scale screening of yeast mutants for sensitvity to the IMP dehydrogenase inhibitor 6-azauracil. Yeast *21*, 241-248.

Robinson,P.J., Trnka,M.J., Pellarin,R., Greenberg,C.H., Bushnell,D.A., Davis,R., Burlingame,A.L., Sali,A., and Kornberg,R.D. (2015). Molecular architecture of the yeast Mediator complex. Elife. *4*.

Roeder,R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem. Sci. *21*, 327-35.

Roguev,A., Wiren,M., Weissman,J.S., and Krogan,N.J. (2007a). High-throughput genetic interaction mapping in the fission yeast Schizosaccharomyces pombe. Nat. Methods *4*, 861-866.

Roguev,A., Wiren,M., Weissman,J.S., and Krogan,N.J. (2007b). High-throughput genetic interaction mapping in the fission yeast Schizosaccharomyces pombe. Nat. Methods *4*, 861-866.

Sadaie,M., Iida,T., Urano,T., and Nakayama,J. (2004). A chromodomain protein, Chp1, is required for the establishment of heterochromatin in fission yeast. EMBO J. *23*, 3825-3835.

Sadeghi,L., Siggens,L., Svensson,J.P., and Ekwall,K. (2014a). Centromeric histone H2B monoubiquitination promotes noncoding transcription and chromatin integrity. Nat. Struct. Mol. Biol. *21*, 236-243.

Sadeghi,L., Siggens,L., Svensson,J.P., and Ekwall,K. (2014b). Centromeric histone H2B monoubiquitination promotes noncoding transcription and chromatin integrity. Nat. Struct. Mol. Biol. *21*, 236-243.

Sanso,M., Lee,K.M., Viladevall,L., Jacques,P.E., Page,V., Nagy,S., Racine,A., St Amour,C.V., Zhang,C., Shokat,K.M., Schwer,B., Robert,F., Fisher,R.P., and Tanny,J.C. (2012a). A positive feedback loop links opposing functions of P-TEFb/Cdk9 and histone H2B ubiquitylation to regulate transcript elongation in fission yeast. PLoS. Genet. *8*, e1002822.

Sanso,M., Lee,K.M., Viladevall,L., Jacques,P.E., Page,V., Nagy,S., Racine,A., St Amour,C.V., Zhang,C., Shokat,K.M., Schwer,B., Robert,F., Fisher,R.P., and Tanny,J.C. (2012b). A positive feedback loop links opposing functions of P-TEFb/Cdk9 and histone H2B ubiquitylation to regulate transcript elongation in fission yeast. PLoS. Genet. *8*, e1002822.

Sato,S., Tomomori-Sato,C., Parmely,T.J., Florens,L., Zybailov,B., Swanson,S.K., Banks,C.A., Jin,J., Cai,Y., Washburn,M.P., Conaway,J.W., and Conaway,R.C. (2004). A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. Mol. Cell *14*, 685-691.

Schwer,B. and Shuman,S. (2011). Deciphering the RNA polymerase II CTD code in fission yeast. Mol. Cell *43*, 311-318.

Shilatifard,A. (1998a). Factors regulating the transcriptional elongation activity of RNA polymerase II. FASEB J. *12*, 1437-1446.

Shilatifard,A. (1998b). Identification and purification of the Holo-ELL complex. J. Biol. Chem. *273*, 11212-11217.

Shilatifard,A., Conaway,J.W., and Conaway,R.C. (1997a). Mechanism and regulation of transcriptional elongation and termination by RNA polymerase II. Curr. Opin. Genet. Dev. *7*, 199-204.

Shilatifard,A., Duan,D.R., Haque,D., Florence,C., Schubach,W.H., Conaway,J.W., and Conaway,R.C. (1997b). ELL2, a new member of an ELL family of RNA polymerase II elongation factors. Proc. Natl. Acad. Sci. U. S. A. *94*, 3639-3643.

Shilatifard,A., Haque,D., Conaway,R.C., and Conaway,J.W. (1997c). Structure and function of RNA polymerase II elongation factor ELL: Identification of two overlapping ELL

functional domains that govern its interaction with polymerase and the ternary elongation complex. J. Biol. Chem. *272*, 22355-22363.

Shilatifard,A., Lane,W.S., Jackson,K.W., Conaway,R.C., and Conaway,J.W. (1996). The human ELL gene encodes a novel RNA polymerase II elongation factor. Science *271*, 1873-1876.

Shinobu,N., Maeda,T., Aso,T., Koike,K., and Hatakeyama,M. (1999). Physical interaction and functional antagonism between the RNA polymerase II elongation factor ELL and p53. J. Biol. Chem. *274*, 17003-17010.

Simone,F., Luo,R.T., Polak,P.E., Kaberlein,J.J., and Thirman,M.J. (2003). ELL-associated factor 2 (EAF2), a functional homolog of EAF1 with alternative ELL binding properties. Blood *101*, 2355-2362.

Simone,F., Polak,P.E., Kaberlein,J.J., Luo,R.T., Levitan,D.A., and Thirman,M.J. (2001). EAF1, a novel ELL-associated factor that is delocalized by expression of the MLL-ELL fusion protein. Blood *98*, 201-209.

Smith,E.R., Lin,C., Garrett,A.S., Thornton,J., Mohaghegh,N., Hu,D., Jackson,J., Saraf,A., Swanson,S.K., Seidel,C., Florens,L., Washburn,M.P., Eissenberg,J.C., and Shilatifard,A. (2011). The little elongation complex regulates small nuclear RNA transcription. Mol. Cell *44*, 954-965.

Sobhian,B., Laguette,N., Yatim,A., Nakamura,M., Levy,Y., Kiernan,R., and Benkirane,M. (2010). HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. Mol. Cell *38*, 439-451.

Spahr,H., Samuelsen,C.O., Baraznenok,V., Ernest,I., Huylebroeck,D., Remacie,J.E., Samuelsson,T., Kieselbach,T., Holmberg,S., and Gustafsson,C.M. (2001). Analysis of Schizosaccharomyces pombe mediator reveals a set of essential subunits conserved between yeast and metazoan cells. Proc. Natl. Acad. Sci. U. S. A. *98*, 11985-11990.

Sugiyama,T., Cam,H.P., Sugiyama,R., Noma,K., Zofall,M., Kobayashi,R., and Grewal,S.I. (2007). SHREC, an effector complex for heterochromatic transcriptional silencing. Cell *128*, 491-504.

Sun,H., Damez-Werno,D.M., Scobie,K.N., Shao,N.Y., Dias,C., Rabkin,J., Koo,J.W., Korb,E., Bagot,R.C., Ahn,F.H., Cahill,M.E., Labonte,B., Mouzon,E., Heller,E.A., Cates,H., Golden,S.A., Gleason,K., Russo,S.J., Andrews,S., Neve,R., Kennedy,P.J., Maze,I., Dietz,D.M., Allis,C.D., Turecki,G., Varga-Weisz,P., Tamminga,C., Shen,L., and Nestler,E.J. (2015). ACF chromatin-remodeling complex mediates stress-induced depressive-like behavior. Nat. Med. *21*, 1146-1153.

Sun,Z.W. and Allis,C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104-108.

Sweta,K., Dabas,P., Jain,K., and Sharma,N. (2017). The amino-terminal domain of ELL transcription elongation factor is essential for ELL function in Schizosaccharomyces pombe. Microbiology *163*, 1641-1653.

Takagi,Y. and Kornberg,R.D. (2006). Mediator as a general transcription factor. J. Biol. Chem. *281*, 80-89.

Takahashi,H., Parmely,T.J., Sato,S., Tomomori-Sato,C., Banks,C.A.S., Kong,S.E., Szutorisz,H., Swanson,S.K., Martin-Brown,S., Washburn,M.P., Florens,L., Seidel,C., Lin,C., Smith,E.R., Shilatifard,A., Conaway,R.C., and Conaway,J.W. (2011). Human Mediator Subunit Med26 Functions As A Docking Site For Transcription Elongation Factors. Cell *146*, 92-104.

Takahashi,H., Takigawa,I., Watanabe,M., Anwar,D., Shibata,M., Tomomori-Sato,C., Sato,S., Ranjan,A., Seidel,C.W., Tsukiyama,T., Mizushima,W., Hayashi,M., Ohkawa,Y., Conaway,J.W., Conaway,R.C., and Hatakeyama,S. (2015). MED26 regulates the transcription of snRNA genes through the recruitment of little elongation complex. Nat. Commun. *6*, 5941.

Thirman,M.J., Diskin,E.B., Bin,S.S., Ip,H.S., Miller,J.M., and Simon,M.C. (1997). Developmental analysis and subcellular localization of the murine homologue of *ELL*. Proc. Natl. Acad. Sci. U. S. A. *94*, 1408-1413.

Thirman,M.J., Levitan,D.A., Kobayashi,H., Simon,M.C., and Rowley,J.D. (1994). Cloning of *ELL*, a gene that fuses to *MLL* in a t(11;19)(q23;p13.1) in acute myeloid leukemia. Proc. Natl. Acad. Sci. U. S. A. *91*, 12110-12114.

Thompson,C.M. and Young,R.A. (1995). General requirement for RNA polymerase II holoenzymes in vivo. Proc. Natl. Acad. Sci. U. S. A *92*, 4587-4590.

Thon,G. and Klar,A.J. (1992). The clr1 locus regulates the expression of the cryptic mating-type loci of fission yeast. Genetics *131*, 287-296.

Tkachuk,D.C., Kohler,S., and Cleary,M.L. (1992). Involvement of a homolog of Drosophila Trithorax by 11q23 chromosomal translocations in acute leukemias. Cell. *71*, 691-700.

Tsai,K.L., Tomomori-Sato,C., Sato,S., Conaway,R.C., Conaway,J.W., and Asturias,F.J. (2014). Subunit architecture and functional modular rearrangements of the transcriptional mediator complex. Cell *157*, 1430-1444.

Tsai,K.L., Yu,X., Gopalan,S., Chao,T.C., Zhang,Y., Florens,L., Washburn,M.P., Murakami,K., Conaway,R.C., Conaway,J.W., and Asturias,F.J. (2017). Mediator structure and rearrangements required for holoenzyme formation. Nature *544*, 196-201.

Uptain,S.M., Kane,C.M., and Chamberlin,M.J. (1997). Basic mechanisms of transcript elongation and its regulation. Annu. Rev. Biochem. *66*, 117-172.

Vedadi,M., Blazer,L., Eram,M.S., Barsyte-Lovejoy,D., Arrowsmith,C.H., and Hajian,T. (2017). Targeting human SET1/MLL family of proteins. Protein Sci. *26*, 662-676.

Wada,T., Takagi,T., Yamaguchi,Y., Ferdous,A., Imai,T., Hirose,S., Sugimoto,S., Yano,K., Hartzog,G.A., Winston,F., Buratowski,S., and Handa,H. (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev. *12*, 343-356.

Wade,P.A., Werel,W., Fentzke,R.C., Thompson,N.E., Leykam,J.F., Burgess,R.R., Jaehning,J.A., and Burton,Z.F. (1996). A novel collection of accessory factors associated with yeast RNA polymerase II. Protein Expr. Purif. *8*, 85-90.

Walmacq,C., Kireeva,M.L., Irvin,J., Nedialkov,Y., Lubkowska,L., Malagon,F., Strathern,J.N., and Kashlev,M. (2009). Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II. J. Biol. Chem. *284*, 19601-19612.

Wang,D., Bushnell,D.A., Huang,X., Westover,K.D., Levitt,M., and Kornberg,R.D. (2009). Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. Science *324*, 1203-1206.

Wang,G., Balamotis,M.A., Stevens,J.L., Yamaguchi,Y., Handa,H., and Berk,A.J. (2005). Mediator requirement for both recruitment and postrecruitment steps in transcription initiation. Mol. Cell *17*, 683-694.

Wang,W., Yao,X., Huang,Y., Hu,X., Liu,R., Hou,D., Chen,R., and Wang,G. (2013). Mediator MED23 regulates basal transcription in vivo via an interaction with P-TEFb. Transcription. *4*, 39-51.

Washburn,M.P., Wolters,D., and Yates,J.R., III (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat. Biotechnol. *19*, 242-247.

Wei,P., Garber,M.E., Fang,S., Fischer,W.H., and Jones,K.A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell *92*, 451-462.

Wolters,D., Washburn,M.P., and Yates,J.R. (2001). An automated multidimensional protein identification technology for shotgun proteomics. Anal. Chem. *73*, 5683-5690.

Yamaguchi,Y., Takagi,T., Wada,T., Yano,K., Furuya,A., Sugimoto,S., Hasegawa,J., and Handa,H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell *97*, 41-51.

Yan,Q., Moreland,R.J., Conaway,J.W., and Conaway,R.C. (1999). Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. J. Biol. Chem. *274*, 35668-75.

Yasukawa,T., Kamura,T., Kitajima,S., Conaway,R.C., Conaway,J.W., and Aso,T. (2008). Mammalian Elongin A complex mediates DNA-damage-induced ubiquitylation and degradation of Rpb1. EMBO J. *27*, 3256-3266.

Yoh,S.M., Lucas,J.S., and Jones,K.A. (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. Genes Dev. *22*, 3422-3434.

Yokoyama,A., Lin,M., Naresh,A., Kitabayashi,I., and Cleary,M.L. (2010). A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. Cancer Cell *17*, 198-212.

Yudkovsky,N., Ranish,J.A., and Hahn,S. (2000). A transcription reinitiation intermediate that is stabilized by activator. Nature *408*, 225-229.

Zhang,K., Mosch,K., Fischle,W., and Grewal,S.I. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat. Struct. Mol. Biol. *15*, 381-388.

Zhou,J., Feng,X., Ban,B., Liu,J., Wang,Z., and Xiao,W. (2009). Elongation factor ELL (Eleven-Nineteen Lysine-rich Leukemia) acts as a transcription factor for direct thrombospondin-1 regulation. J. Biol. Chem. *284*, 19142-19152.

Zhou,Q., Li,T., and Price,D.H. (2012). RNA polymerase II elongation control. Annu. Rev. Biochem. *81*, 119-143.

Zhu,X., Wiren,M., Sinha,I., Rasmussen,N.N., Linder,T., Holmberg,S., Ekwall,K., and Gustafsson,C.M. (2006). Genome-wide occupancy profile of mediator and the Srb8-11 module reveals interactions with coding regions. Mol. Cell *22*, 169-178.